

GenCore version 5.1.16  
Copyright (C) 1993 - 2003 Computer Ltd.

OM protein protein search, using sw model

Run on: May 30, 2003, 15:55:12, Search time 9.08708 seconds  
(without alignments)  
362.883 Million cell updates/sec

Title: US-09-872-185b-1  
Perfect score: 595  
Sequence: 1 AONTIARIKEPILVIRKAGAP... VPPVQBACKPQIVISASFLI 112

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued\_Patents\_AA: \*  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/1/1aa/5H\_COMB.pep.\*  
3: /cgn2\_6/prodata/1/1aa/5A\_3MB.pep.\*  
4: /cgn2\_6/prodata/1/1aa/6R\_COMB.pep.\*  
5: /cgn2\_6/prodata/1/1aa/6CTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/1/1aa/backfiles1.pep.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	595	100.0	418	2	US-08-633-148-4
2	595	100.0	342	4	US-09-062-365-1
3	595	100.0	340	2	US-08-533-148-2
4	558	93.8	278	2	US-08-432-016-5
5	558	93.8	278	2	US-08-684-594-5
6	162	27.2	30	4	US-04-062-365-5
7	114	19.2	144	4	US-09-341-886-25
8	114	19.2	144	5	PCI US91 05277-2
9	97	16.3	607	2	US-08-752-307B-12
10	97	16.3	607	4	US-09-707-802-12
11	97	16.3	607	4	US-09-991-326-13
12	96	16.3	124	4	US-08-506-246B-14
13	96	16.1	96	2	US-08-341-843B-31
14	96	16.1	96	2	US-08-427-497E-16
15	95.5	16.1	611	2	US-08-752-307B-10
16	95.5	16.1	611	4	US-09-707-802-10
17	95.5	16.1	611	4	US-09-991-326-10
18	93	15.6	804	4	US-09-877-730-4
19	93	15.6	804	4	US-09-877-730-6
20	93	15.6	965	4	US-09-877-730-10
21	93	15.6	1078	1	US-08-408-091-6
22	93	15.6	1018	1	US-08-408-420A-6
23	93	15.6	1018	1	US-08-714-991-6
24	93	15.6	1018	3	US-08-040-741-6
25	93	15.6	1069	4	US-09-877-730-2
26	93	15.6	1150	4	US-09-877-730-8
27	92.5	15.5	1651	4	US-09-540-245A-18

28 92 15.5 642 1 US-08-237-299-1  
29 92 15.5 698 2 US-08-602-725-36  
30 92 15.5 734 2 US-08-389-459A-17  
31 92 15.5 734 3 US-08-987-867A-17  
32 91 15.3 466 2 US-08-432-016-4  
33 91 15.3 466 2 US-08-684-594-4  
34 87 14.6 431 4 US-09-038-832-2  
35 87 14.6 431 4 US-09-038-832-4  
36 86.5 14.5 91 2 US-08-341-843B-31  
37 86.5 14.5 91 2 US-08-427-497E-16  
38 86.5 14.5 1266 4 US-08-506-246B-4  
39 86 14.5 421 2 US-08-659-984A-1  
40 86 14.5 421 4 US-08-660-531-1  
41 86 14.5 444 2 US-08-659-984A-5  
42 86 14.5 444 4 US-08-660-531-5  
43 86 14.5 605 2 US-08-752-307B-8  
44 86 14.5 605 4 US-09-707-802-8  
45 86 14.5 605 4 US-09-991-326-8

ALIGNMENTS

RESULT 1  
US-08-633-148-4  
Sequence 4, Application US/08633148  
Patent No. 5864018  
GENERAL INFORMATION:  
APPLICANT: MORSE, MICHAEL J.  
APPLICANT: NAGASHIMA, MARIKO  
TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION  
TITLE OF INVENTION: END PEPTIDE REPEAT POLYPEPTIDES AND USES THEREFOR  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND & TOWNSEND & CREW LLP  
STREET: TWO EMBARCADERO CENTER, 8TH FLOOR  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: U.S.A.  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/633,148  
FILING DATE: 16-APR-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURPHY ESQ., MATTHEW B.  
REGISTRATION NUMBER: 19,787  
ADDRESS/ADDRESS NUMBER: 014618-00560005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQUENCE NO. 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 318 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-633-148-4

Query Match: 100.0%, Score 595, DB 2, Length 318;  
Best Local Similarity: 100.0%, Prod. No. 8 60-64;  
Matches: 112; Conserved: 0; Mismatches: 0; Indels: 0; Gaps: 0.

QY 1 AONTIARIKEPILVIRKAGAPKPPQIVISASFLI 60  
|||||  
LB 1 AONTIARIKEPILVIRKAGAPKPPQIVISASFLI 60



Best Local Similarity 99.08, Pred. No. 2, 1e-59,  
Matches 104; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 IGEPLVLKCKGAPKKPQRLNIGRTIPANKVLSPPGSGPWSVARVLPNGSLFLPAV 67  
Db 1 IGEPLVLKCKGAPKKPQRLNIGRTIPANKVLSPPGSGPWSVARVLPNGSLFLPAV 60  
QY 68 GIQDEGIFRCAMRNNGKETKSNYRVVYQIPCKPEIVDSASLT 112  
Db 61 GIQDEGIFRCAMRNNGKETKSNYRVVYQIPCKPEIVDSASLT 105

## RESULT 5

US-08-684-594-5  
Sequence 5, Application US/08684594

Patent No. 5998172

GENERAL INFORMATION:

APPLICANT: HAYNES, BARTON E.

APPLICANT: ARUFFO, ALFONSO

APPLICANT: PATEL, DHAVAKUMAR

APPLICANT: HOWEN, MICHAEL A.

APPLICANT: MARQUARDT, HANS

TITLE OF INVENTION: CD6 LIGAND

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHIVE, P.C.

STREET: 1100 NORTH CLIFTON ROAD

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/084,594

FILING DATE: 18-JUL-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/432,016

FILING DATE: 01-MAY-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/333,350

FILING DATE: 02-NOV-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/143,903

FILING DATE: 02-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: WILSON, MARY J.

REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 1579-112

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4900

TELEFAX: (703) 816-4100

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 278 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-684-594-5

Query Match

Best Local Similarity 99.08, Pred. No. 2, 1e-59, Length 278,

Matches 104; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 IGEPLVLKCKGAPKKPQRLNIGRTIPANKVLSPPGSGPWSVARVLPNGSLFLPAV 67

Db 1 IGEPLVLKCKGAPKKPQRLNIGRTIPANKVLSPPGSGPWSVARVLPNGSLFLPAV 60

QY 68 GIQDEGIFRCAMRNNGKETKSNYRVVYQIPCKPEIVDSASLT 112  
Db 61 GIQDEGIFRCAMRNNGKETKSNYRVVYQIPCKPEIVDSASLT 105

## RESULT 6

US-09-062-365-5

Sequence 5, Application US/99042365

Patent No. 6465422

GENERAL INFORMATION:

APPLICANT: Schmidt, Ann Marie

APPLICANT: Stern, David

TITLE OF INVENTION: METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A

TITLE OF INVENTION: SUBJECT

FILE REFERENCE: 55424

CURRENT APPLICATION NUMBER: US/09/062,365

CURRENT FILING DATE: 1998-04-17

NUMBER OF SEQ ID NOS: 6

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 5

LENGTH: 30

TYPE: PRT

ORGANISM: Human

US-09-062-365-5

Query Match

Best Local Similarity 100.0%; Pred. No. 6, 3e-13; Length 30,

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQNTAIGEPVLKCKGAPKKPQRLNIGRTIPANKVLSPPGSGPWSVARVLPNGSLFLPAV 30

Db 1 AQNTAIGEPVLKCKGAPKKPQRLNIGRTIPANKVLSPPGSGPWSVARVLPNGSLFLPAV 30

## RESULT 7

US-09-041-886-25

Sequence 25, Application US/09041886

Patent No. 6235872

GENERAL INFORMATION:

APPLICANT: Bredesen, Dale E.

APPLICANT: Rabinzadeh, Shantoz

TITLE OF INVENTION: Erythropoietic Peptides, Dependence

TITLE OF INVENTION: Polypeptides and Methods of Use

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: Campbell & Flores LLP

STREET: 4370 La Jolla Village drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: United States

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/041,886

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-IJ 2626

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:

LENGTH: 1447 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

us-09-872-185b-1.rai

Fri May 30 17:16:31 2003

```

US-09 041-886-25
Query Match 19.2% Score 114; DB 4; Length 1447;
Best Local Similarity 41.4% Pred. No. 8e-05;
Matches 41; Conservative 20; Mismatches 38; Indels 10; Gaps 3;

QY 2 QNITAKGEPLVLCGKAPKKPPQRLFKLNTGRTFAWKVLSPGCGGWDSVARVLPNGS 61
DB 147 ESVTAFMGDTVLKCGE-VIGEPMTTHWQKQ-----QDLTPI---PCDSRVVVLPSGA 196

QY 62 LFLAVGIGQDEIFECQAMNKNKETSNTYRVVYVLPNG 100
DB 197 LQISRLQPGDIGIVRCASRNPAASSRTGNEAEVRILSDPG 245

RESULT 8
PCT-US94-05277.2
Sequence 2, Application US/US940405277
GENERAL INFORMATION:
APPLICANT: Bruskin, Arthur
APPLICANT: Jarosz, David E.
APPLICANT: Johnson, Karen
APPLICANT: Kinzler, Kenneth W.
APPLICANT: Vogelstein, Bert
APPLICANT: Zaretsky, James R.
TITLE OF INVENTION: Antibodies Specific for DCC Gene Product
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05277
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 42,141
REFERENCE/DOCKET NUMBER: 01107.42709
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9294
TELEX: 197430 BMB UT
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1447 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-05277.2

Query Match 19.2% Score 114; DB 5; Length 1447;
Best Local Similarity 41.4% Pred. No. 8e-05;
Matches 41; Conservative 20; Mismatches 38; Indels 10; Gaps 3;

QY 2 QNITAKGEPLVLCGKAPKKPPQRLFKLNTGRTFAWKVLSPGCGGWDSVARVLPNGS 61
DB 147 ESVTAFMGDTVLKCGE-VIGEPMTTHWQKQ-----QDLTPI---PCDSRVVVLPSGA 196

QY 62 LFLAVGIGQDEIFECQAMNKNKETSNTYRVVYVLPNG 100
DB 197 LQISRLQPGDIGIVRCASRNPAASSRTGNEAEVRILSDPG 245

RESULT 9
PCT-US94-05277.2
Sequence 2, Application US/US940405277
GENERAL INFORMATION:
APPLICANT: Bruskin, Arthur
APPLICANT: Jarosz, David E.
APPLICANT: Johnson, Karen
APPLICANT: Kinzler, Kenneth W.
APPLICANT: Vogelstein, Bert
APPLICANT: Zaretsky, James R.
TITLE OF INVENTION: Antibodies Specific for DCC Gene Product
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: IBM PC compatible
OPERATING SYSTEM: PC DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05277
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 42,141
REFERENCE/DOCKET NUMBER: 01107.42709
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9294
TELEX: 197430 BMB UT
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1447 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-05277.2

Query Match 16.3% Score 97; DB 2; Length 607;
Best Local Similarity 40.5% Pred. No. 0.0027;
Matches 25; Conservative 16; Mismatches 25; Indels 16; Gaps 4;

QY 6 ARICGPIVLKCGKAPKKPPQRLFKLNTGRTFAWKVLSPGCGGWDSVARVLPNGSLELP 65
DB 251 ALVGGQVTLGEC-FAFGNEVPRIKRWKVDGS-----LSPG WTTA EPTLOIP 294

QY 66 AVGIQDEIFECQAMNKNKETSNTYRVVYVLPNG 87
DB 295 SVSFEDEGTYECFAFNSKGRPI 416

RESULT 10
US-09 707-802-12
Sequence 12, Application US/09707802
Patent No. 6391586
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
APPLICANT: Gearing, David P.
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
ENCODING NOVEL SECRETED OR MEMBRANE ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston

```



```

?
? STATE: MA
? COUNTRY: US
? ZIP: 02110-2804
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: Windows95
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/707,802
? FILING DATE: 07-NOV-1996
? CLASSIFICATION: <Unknown>
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/752,307
? FILING DATE: <Unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Meiklejohn, Ph.D., Anita L.
? REFERENCE/DOCKET NUMBER: 35,283
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-542-5070
? TELEFAX: 617-542-8906
? TELEX: 200154
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 607 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-707-802-12

Query Match 16.3%, Score 97, DB 4; Length 607;
Best Local Similarity 30.5%, Pred. No. 0.0027;
Matches 25; Conservative 16; Mismatches 25; Indels 16; Gaps 4;

QY 6 ARICEP-VLKCKGA-KKPKGLEWKNIGKTEAKKVLSEGGSPWUSVAKVLPNSLSLFLP 65
DB 251 ALVQQVTLECFAGNPVPRIKRWKVDG-----LSQ-----WTTA-----EPTLOIP 294

QY 66 AVGIQDGGIFPCQAMNNGKET 87
DB 295 SVSFEDEGTVECAENSKGEDI 316

RESULT 11
US-09-991-326-12
? Sequence 12, Application US/09091326
? Patent No. 6395872
? GENERAL INFORMATION:
? APPLICANT: McCarthy, Sean A.
? CO-INVENTOR: David P.
? CO-INVENTOR: Douglas A.
? TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
? ENCODING NAVAL SECRETED UP MEMBRANE-ASSOCIATED PROTEIN
? NUMBER OF SEQUENCES: 14
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Fish & Richardson, P.C.
? STREET: 225 Franklin Street
? CITY: Boston
? STATE: MA
? COUNTRY: US
? ZIP: 02110-2804
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: Windows95
? SOFTWARE: FASTSEQ for Windows Version 2.0
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/991,326
? FILING DATE: 21-NOV-1996
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/752,307

```

```

?
? FILING DATE: 19-NOV-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Meiklejohn, Ph.D., Anita L.
? REGISTRATION NUMBER: 35,283
? REFERENCE/DOCKET NUMBER: 09404/020002
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-542-5070
? TELEFAX: 617-542-8906
? TELEX: 200154
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 607 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-991-326-12

Query Match 16.3%, Score 97, DB 4; Length 607;
Best Local Similarity 30.5%, Pred. No. 0.0027;
Matches 25; Conservative 16; Mismatches 25; Indels 16; Gaps 4;

QY 6 ARICEP-VLKCKGA-KKPKGLEWKNIGKTEAKKVLSEGGSPWUSVAKVLPNSLSLFLP 65
DB 251 ALVQQVTLECFAGNPVPRIKRWKVDG-----LSQ-----WTTA-----EPTLOIP 294

QY 66 AVGIQDGGIFPCQAMNNGKET 87
DB 295 SVSFEDEGTVECAENSKGEDI 316

RESULT 12
US-08-506-296B-14
? Sequence 14, Application US/08506296B
? Patent No. 6313265
? GENERAL INFORMATION:
? APPLICANT: Williams, Greg
? APPLICANT: Cunningham, Bruce A.
? APPLICANT: Crossin, Kathryn L.
? TITLE OF INVENTION: NEURITE OUTGROWTH-PROMOTING POLYPEPTIDES
? TITLE OF INVENTION: CONTAINING FIROHNETIN TYPE III REPEATS AND METHODS OF USE
? NUMBER OF SEQUENCES: 77
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: The Scripps Research Institute
? STREET: 10550 No. 6313265th Torrey Pines Road, TPC-8
? CITY: La Jolla
? STATE: California
? COUNTRY: U.S.
? ZIP: 92037
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/506,296B
? FILING DATE: 24-JUL-1995
? CLASSIFICATION: 514
? ATTORNEY/AGENT INFORMATION:
? NAME: Filling, Thomas
? REGISTRATION NUMBER: 34,153
? REFERENCE/DOCKET NUMBER: TSRI 488.0
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619) 554-2937
? TELEFAX: (619) 554-6312
? INFORMATION FOR SEQ ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1253 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
US-08-506-296B-14

```



```

; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; INDIVIDUAL ISOLATE: 17-18 week fetus
; IMMEDIATE SOURCE:
; LIBRARY: Stratagene cDNA library 936206
; CLONE: synthesis of 4 clones
; PUBLICATION INFORMATION:
; AUTHORS: Ilavina, Mary Louise
; CO-AUTHORS: Lemmon, Vance
; TITLE: Molecular structure and functional
; TITLE: testing of human IL6M: an interspecies
; TITLE: comparison.
; JOURNAL: GENOMICS
; VOLUME: 11
; ISSUE:
; PAGES: 416-423
; DATE: 1991
; US-08-427-497P-23

Query Match: 16.1%; Score 96; DB 4; Length 96;
Best Local Similarity 42.7%; Pred. No. 0.00029;
Matches 33; Conservative 11; Mismatches 39; Indels 18; Gaps 3;

QY 1 AONITAHIGPIVLKC--KGAPKPPQRLKWLKLTGRTKAWKVLSPGGGWDVSVARVLP 58
Db 12 SSHLVALQGQPLVLECIAGFP .....TPTIKWLRLPSGMPADRVTYQNH 56
QY 59 NGSILFLPVGIGDEGIFRCQAMNRNGKETKSNRYRVYQIP 99
Db 57 NKTILQLLKVGEDDDGEYRCIAENSLG-SARHAYYVTVLAAP 96

RESULT 15
US-08-752-307B-10
Sequence 10; Application US/08752307B
Patent No. 5952171
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
APPLICANT: Gearing, David P.
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
NUMBER OF INVENTIONS: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752.307B
FILING DATE: 19-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejoh, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: U9404/020001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 611 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-752-307B-10

Query Match: 16.1%; Score 95.5; DB 2; Length 611;
Best Local Similarity 32.3%; Pred. No. 0.0042;
Matches 30; Conservative 11; Mismatches 35; Indels 17; Gaps 2;

QY 1 AONITAHIGPIVLKC--KGAPKPPQRLKWLKLTGRTKAWKVLSPGGGWDVSVARVLP 58
Db 249 SSHLVALQGQPLVLECIAGFP .....TPTIKWLRLPSGMPADRVTYQNH 293
QY 59 NGSILFLPVGIGDEGIFRCQAMNRNGKETKSNY 91
Db 294 NKTILQLLKVGEDDDGEYRCIAENSLGSAHAYY 326

Search completed: May 30, 2003, 15.59.50
Job time: 10.0811 secs

```





PT Detecting a receptor for advanced glycosylated end-products (RAGE)  
 PT modulators, for treating e.g., cancer, diabetes or inflammation,  
 PT comprises measuring the amount of bound anti-RAGE antibody  
 XX  
 PS Claim 4; Fig 2; 49pp; English.  
 XX  
 CC The invention relates to detection receptor for advanced glycosylated  
 CC end-products (PAGE) modulators comprising determining the amount of PAGE  
 CC protein or its fragment bound to the pre-adsorbed liquid by measuring the  
 CC amount of anti-RAGE antibody bound to the solid surface, the method is  
 CC useful for rapid, high-throughput identification of compounds that  
 CC modulate RAGE. The compounds are useful for treating symptoms of diabetes  
 CC and symptoms of diabetic late complications, amyloidosis, Alzheimer's  
 CC disease, cancer, inflammation, kidney failure, systemic lupus nephritis  
 CC or inflammatory lupus nephritis, erectile dysfunction and  
 CC atherosclerosis.  
 XX  
 CC Sequence 112 AA:  
 SQ Query Match 100.0%; Score 595; DB 24; Length 112;  
 Best Local Similarity 100.0%; Pred. No. 5; 96;  
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AUNTARIGEPVLVKCKAPKKPQRLWKLNTGRTAEWKVLSPPGGGPDWSVARVLPNG 60  
 DB 1 AUNTARIGEPVLVKCKAPKKPQRLWKLNTGRTAEWKVLSPPGGGPDWSVARVLPNG 60  
 QY 61 SLFLPAVGLODEGIFRCQAMNPNCKETKSNRYRVVYQIPCKPELVDSASELT 112  
 DB 61 SLFLPAVGLODEGIFRCQAMNPNCKETKSNRYRVVYQIPCKPELVDSASELT 112  
 RESULT 2  
 AAW44200  
 ID AAW44200 standard; Protein: 418 AA.  
 AC AAW44200;  
 XX  
 XX 14 MAY 1998 (first entry)  
 DI Human mature receptor to an advanced glycosylation end product.  
 DE Human; soluble receptor; advanced glycosylation end product; RAGE;  
 KW AGE; antibody; vascular permeability; diabetes mellitus.  
 KW Homo sapiens.  
 CS  
 XX Key Location/Qualifiers  
 FH Misc difference 66  
 FT /note- "encoded by cct"  
 XX  
 UN W0979125-A1.  
 XX  
 XX 24 OCT 1997.  
 PD 11 APR 1997; 97WO EP01844.  
 PF 16 APR 1996; 96US 0633146.  
 XX  
 XX (SCH) SCHERING PATENT AG;  
 FA Hollander DA, Morser MJ, Nagashima M;  
 FI WPI: 1997 526458/51.  
 DE N PSDB: AAV12495.  
 XX  
 XX Anti advanced glycosylation end product polypeptide antibody  
 PT prevents receptor binding and therefore reduces vascular  
 PT permeability, useful to treat diabetes mellitus  
 XX  
 XX Claim 2; Page 42-44; 9pp; English.  
 PS  
 XX The present sequence represents a mature human receptor to an advanced

CC glycosylation end product (RAGE) polypeptide. The present invention  
 CC describes an isolated antibody (Ab), specifically immunoreactive with  
 CC RAGE. Advanced glycosylation end products (AGE) of proteins are  
 CC non-enzymatically glycosylated proteins, which accumulate in vascular  
 CC tissue in aging, and at an accelerated rate in individuals with  
 CC diabetes. The Ab, which prevents the interaction between AGE and its  
 CC receptor (RAGE), reduces vascular permeability. The Ab can be used to  
 CC treat diabetes mellitus symptoms, e.g., microvascular pathology, occlusive  
 CC vascular disorders, neuropathy, nephropathy, retinopathy, haemodialysis  
 CC associated amyloidosis or atherosclerosis. The Ab can also be used for  
 CC the isolation and purification of human RAGE polypeptide.  
 XX  
 CC Sequence 318 AA:  
 SQ Query Match 100.0%; Score 595; DB 18; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 26; 55;  
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AUNTARIGEPVLVKCKAPKKPQRLWKLNTGRTAEWKVLSPPGGGPDWSVARVLPNG 60  
 DB 1 AUNTARIGEPVLVKCKAPKKPQRLWKLNTGRTAEWKVLSPPGGGPDWSVARVLPNG 60  
 QY 61 SLFLPAVGLODEGIFRCQAMNPNCKETKSNRYRVVYQIPCKPELVDSASELT 112  
 DB 61 SLFLPAVGLODEGIFRCQAMNPNCKETKSNRYRVVYQIPCKPELVDSASELT 112  
 RESULT 4  
 AAW44754  
 ID AAW44754 standard; Protein: 418 AA.  
 AC AAW44754;  
 XX  
 XX 08 MAY 1998 (first entry)  
 DI Human RAGE polypeptide (418 amino acid residues).  
 DE Advanced glycosylation end-product receptor; RAGE; serpinin; AGE;  
 KW vascular permeability; diabetes mellitus; treatment; atherosclerosis;  
 KW Alzheimer's disease.  
 XX  
 XX Homo sapiens.  
 XX W0979121-A1.  
 FN 24 OCT 1997.  
 PD 11 APR 1997; 97WO EP01844.  
 PF 16 APR 1996; 96US 0633147.  
 XX  
 XX (SCH) SCHERING AG.  
 FA Morser MJ, Nagashima M;  
 FI WPI: 1997 526458/48.  
 DE N PSDB: AAV06518.  
 XX  
 XX New soluble advanced glycosylation end product receptor polypeptide  
 PT used for reducing vascular permeability, complications of diabetes  
 PT etc., also for purification and to screen for modulators  
 XX  
 XX Claim 4; Fig 1b; 9pp; English.  
 PS  
 XX This is a human advanced glycosylation end product receptor (PAGE)  
 CC polypeptide (418 amino acid residues). The PAGE polypeptides and its  
 CC active fragments or their mimetics, inhibit interaction between advanced  
 CC glycosylation end products (AGE) and a receptor (specifically RAGE). They  
 CC are used to treat diseases associated with AGE/RAGE interaction, such as  
 CC increased vascular permeability, diabetes mellitus (particularly  
 CC complications such as micro or macro vasculopathy or occlusive vascular  
 CC disorders such as neuropathy, nephropathy, retinopathy or  
 CC atherosclerosis) or haemodialysis associated amyloidosis, also activation

CC of microglial cells by beta amyloid peptides in Alzheimer's disease of  
 CC age related disorders such as oxidative stress. These RAGE polypeptides  
 CC are also used, when immobilised, to purify AGE from a protein mixture and  
 CC to screen for compounds that are agonists and antagonists of AGE/RAGE  
 CC interaction. They can also be used diagnostically to detect abnormal  
 CC levels of AGE. Antibodies against RAGE polypeptides are useful as  
 CC immunosay reagents for measurement of RAGE levels, and as inhibitors of  
 CC interaction between AGE and RAGE or other receptors and for purification  
 CC and quantification of RAGE polypeptides. The following nucleic acids are  
 CC used to express recombinant RAGE and as probes for isolating related  
 CC genes.

XX Sequence 318 AA;  
 XX Query Match 100.0%; Score 535; DB 18; Length 318;  
 XX Best Local Similarity 100.0%; Pred. No 2e-55;  
 XX Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQNITARIIGEPVLKCKGAPKPPQPLEWKLNTGRTFAWKVLSPOGGPWSVARVPLNG 60  
 DB 1 AQNITARIIGEPVLKCKGAPKPPQPLEWKLNTGRTFAWKVLSPOGGPWSVARVPLNG 60  
 QY 61 SLFLPAVGIGEGIFRQAMNNGKETKSNYRVPVYQIPGKPEIVDSASELT 112  
 DB 61 SLFLPAVGIGEGIFRQAMNNGKETKSNYRVPVYQIPGKPEIVDSASELT 112

## RESULT 4

AAV52130  
 ID AAV52130 standard; protein: 332 AA.

AC AAV52130;

DT 28-JAN-2000 (first entry)

XX Human Receptor to AGE (RAGE) amino acid sequence.

XX Soluble receptor for advanced glycation endproducts (AGE) tumour  
 KW invasion; metastasis; amphotericin; neuron; inhibit; therapy.

XX Homo sapiens.

XX W09554485-A1.

XX 28-OCT-1999

XX 16-APR-1999; 99NW-0508427

XX 17-APR-1998; 98US-0062365

XX (UNCO) UNIV COLUMBIA NEW YORK

XX Schmidt AM, Stern D;

XX WPI; 2000-014260/01.

XX Inhibiting tumour invasion or spreading by administration of soluble  
 PT receptor for advanced glycation endproducts -

XX Disclosure; Page 10-11, 88pp, English.

XX This is the amino acid sequence of the human soluble receptor for  
 CC Advanced Glycation Endproducts (AGE). RAGE interacts with a range of  
 CC physiologically and pathophysiologically relevant ligands when  
 CC considering tumour invasion. In normal developing neurons RAGE  
 CC colocalizes with amphoterin which is a matrix associated polypeptide.  
 CC The expression of both RAGE and amphoterin decreases after birth, but  
 CC both have increased expression in tumours. RAGE polypeptides  
 CC AAV52132-Y52135 are used in the invention in a method for inhibiting  
 CC tumour invasion and metastasis. The method involves inhibiting tumour  
 CC invasion and metastasis via administration of a therapeutically effective  
 CC amount of the pharmaceutical composition containing a RAGE polypeptide.  
 CC The invention also relates to a method for evaluating the ability of an

CC agent to inhibit tumour invasion in a local cellular environment. RAGE  
 CC can be administered to a patient in a pharmaceutical y acceptable  
 CC carrier.

XX Sequence 332 AA;

XX Query Match 100.0%; Score 535; DB 21; Length 332;

XX Best Local Similarity 100.0%; Pred. No 2.2e-55;

XX Matches 112; Conservative 0; Mismatches 0; Indels 5; Gaps 0;

QY 1 AQNITARIIGEPVLKCKGAPKPPQPLEWKLNTGRTFAWKVLSPOGGPWSVARVPLNG 60  
 DB 1 AQNITARIIGEPVLKCKGAPKPPQPLEWKLNTGRTFAWKVLSPOGGPWSVARVPLNG 60

QY 61 SLFLPAVGIGEGIFRQAMNNGKETKSNYRVPVYQIPGKPEIVDSASELT 112  
 DB 61 SLFLPAVGIGEGIFRQAMNNGKETKSNYRVPVYQIPGKPEIVDSASELT 112

## RESULT 5

AA48746  
 ID AA48746 standard; protein: 339 AA.

XX AA48746;

XX 02-APR-2002 (first entry)

XX Human SPAGE protein SEQ ID NO 2

XX Human, RAGE, receptor for advanced glycation endproduct, receptor,  
 KW antidiabetic, neuroprotective, cytostatic, antiinflammatory, vasotrophic,  
 KW nephrotropic, dermatologic, antiarteriosclerotic, neurotropic, diabetes;  
 KW Alzheimer's disease; cancer; inflammation; kidney failure;  
 KW systemic lupus, nephritis, erectile dysfunction, atherosclerosis.

XX Homo sapiens.

XX W0200192592-A2.

XX 06-DEC-2001.

XX 30-MAY-2001; 2001WO-US17447.

XX 30-MAY-2000; 2000US-297342P.

XX 05-MAR-2001; 2001US-0799152

XX (TEAN-) TRANS TECH PHARMA.

XX Shahbar M;

XX WPI; 2002-114372/15.

XX Detecting a receptor for advanced glycation endproducts (RAGE)  
 PT modulators, for treating e.g., cancer, diabetes or inflammation,  
 PT comprises measuring the amount of bound anti-RAGE antibody -

XX Claim 2; Fig 2; 49pp; English.

XX The invention relates to detecting receptor for advanced glycation  
 CC endproducts (AGE). modulators comprises determining the amount of RAGE  
 CC protein or its fragment bound to the pre-adsorbed ligand by measuring the  
 CC amount of anti-RAGE antibody bound to the solid surface. The method is  
 CC useful for rapid, high-throughput identification of compounds that  
 CC modulate RAGE. The compounds are useful for treating symptoms of diabetes  
 CC and symptoms of diabetic late complications, amyloidosis, Alzheimer's  
 CC disease, cancer, inflammation, kidney failure, systemic lupus nephritis  
 CC or inflammatory lupus nephritis, erectile dysfunction and  
 CC atherosclerosis.

XX Sequence 339 AA;

XX Query Match 100.0%;

XX Best Local Similarity 100.0%; Score 535; DB 23; Length 339;

XX Matches 112; Conservative 0; Mismatches 0; Indels 5; Gaps 0;





RESULT 8  
AAH81925  
ID AAH81925 standard, protein, 404 AA.  
XX AC AAB81925;  
XX DT 15-JUN-2001 (first entry)  
XX DE Extracorporeal circulation material receptor protein.  
XX KW Extracorporeal circulation; carbonyl stress product; receptor;  
XX KW diabetes; vascular lesion; excretory dysfunction  
XX OS Unidentified.  
XX PN W0200118060-A1.  
XX PD 15-MAR-2001.  
XX PF 08-SEP-1999, 2000W03-0106172.  
XX PR 08-SEP-1999; 99JP-0254463.  
XX PA (TORA) TOPAY IND INC.  
XX PI Shimizu S, Kubota M, Akiyama H, Usui M;  
XX WP1: 2001-290314/30.  
XX Material for extracorporeal circulation, applicable in selective  
PT elimination of diabetic complication factors such as carbonyl stress  
PT products caused by abnormally promoted carbonyl stress from excretory  
PT dysfunction in vascular lesions  
XX claim 1, Page 31-32; 36pp; Japanese.  
XX the present invention describes a material for extracorporeal circulation  
XX which is made from a water-insoluble carrier immobilized with a protein  
CC having the sequence shown here. The materials of the invention, including  
CC adsorbents, are for extracorporeal circulation, which are applicable in  
CC the selective elimination of diabetic complication factors from a body  
CC fluid, and are therefore useful in treating vascular lesions like  
CC arteriosclerosis due to carbonyl stress products caused by abnormally  
CC promoted carbonyl stress from excretory dysfunction.  
XX SQ Sequence 404 AA;  
Query Match 100.0%; Score 595; DB 22; Length 404;  
Best Local Similarity 100.0%; Pred. No. 2.7e-55;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AONTARIGEPVLVKCKGAPKPPORLEWKINTGRTEAWKVLSPQGGPWDSVARVLPNG 60  
Db 23 AONTARIGEPVLVKCKGAPKPPORLEWKINTGRTEAWKVLSPQGGPWDSVARVLPNG 82  
QY 61 SLFLPAVGIDQEGIFRCQAMNRNGKETKSNYRVRYVQIPCKPEIVDSASELT 112  
Db 83 SLFLPAVGIDQEGIFRCQAMNRNGKETKSNYRVRYVQIPCKPEIVDSASELT 134  
RESULT 9  
AAM48745  
ID AAM48745 standard; protein; 404 AA.  
XX AC AAM48745;  
XX DT 02-APR-2002 (first entry)  
XX DE Human RAGE protein SEQ ID NO 1.  
XX KW Human; RAGE; receptor for advanced glycosylated endproduct; receptor;

KW antidiabetic; neuroprotective; cylostatic; antiinflammatory; vasotropic;  
KW nephrotropic; dermatological; antiarteriosclerotic; nootropic; diabetes;  
KW Alzheimer's disease; cancer; inflammation; kidney failure;  
KW systemic lupus; nephritis; erectile dysfunction; atherosclerosis.  
OS Homo sapiens.  
XX PN W0200192892-A2.  
XX PD 06-DEC-2001  
XX PF 30-MAY-2001; 2001W0-US17447.  
XX PP 30-MAY-2000; 2000US-307342P.  
XX PR 05-MAR-2001; 2001US-0799152.  
XX PA (TRAN-) TRANS TECH PHARMA.  
XX PI Shahbaz M;  
XX WP1: 2002-114372/15.  
XX Detecting a receptor for advanced glycosylated endproducts (RAGE)  
PT modulators, for treating e.g., cancer, diabetes or inflammation,  
PT comprises measuring the amount of bound anti-PAGE antibody  
XX Claim 1; Fig 2; 49pp; English.  
XX The invention relates to detecting receptor for advanced glycosylated  
CC endproducts (RAGE) modulators comprising determining the amount of RAGE  
CC protein or its fragment bound to the pre-adsorbed ligand by measuring the  
CC amount of anti-RAGE antibody bound to the solid surface. The method is  
CC useful for rapid, high-throughput identification of compounds that  
CC modulate RAGE. The compounds are useful for treating symptoms of diabetes  
CC and symptoms of diabetic late complications, amyloidosis, Alzheimer's  
CC disease, cancer, inflammation, kidney failure, systemic lupus nephritis  
CC or inflammatory lupus nephritis, erectile dysfunction and  
XX atherosclerosis.  
XX SQ Sequence 404 AA;  
Query Match 100.0%; Score 595; DB 23; Length 404;  
Best Local Similarity 100.0%; Pred. No. 2.7e-55;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AONTARIGEPVLVKCKGAPKPPORLEWKINTGRTEAWKVLSPQGGPWDSVARVLPNG 60  
Db 23 AONTARIGEPVLVKCKGAPKPPORLEWKINTGRTEAWKVLSPQGGPWDSVARVLPNG 82  
QY 61 SLFLPAVGIDQEGIFRCQAMNRNGKETKSNYRVRYVQIPCKPEIVDSASELT 112  
Db 83 SLFLPAVGIDQEGIFRCQAMNRNGKETKSNYRVRYVQIPCKPEIVDSASELT 134  
RESULT 10  
AAE23219  
ID AAE23219 standard; protein; 404 AA.  
XX AC AAE23219;  
XX DT 27-AUG-2002 (first entry)  
XX DE Human receptor for advanced glycosylation end product (RAGE) protein.  
XX KW Human; Receptor for advanced glycosylation end product; RAGE; cardiant;  
KW tissue growth; neointimal formation; blood vessel; restenosis; diabetes;  
KW myocardial infarction; angioplasty; peripheral vascular surgery; angina;  
KW transgenic animal; acute thrombotic stroke; venous thrombosis.  
OS Homo sapiens.  
XX PN W0200250889-A3.  
XX



XX 13-OCT-2000; 2000US-0687528.  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX Stern DM, Schmidt A, Marso S, Topol E, Lincoff AM;  
 XX WPI: 2002-426260/45.  
 XX N-PSDB: AAD36951.  
 XX Inhibiting new tissue growth or neointimal formation in blood vessels  
 PT of subject suffering from diabetes, stroke and preventing restenosis,  
 PT comprises administering inhibitor of receptor for advanced glycation end  
 PT product  
 XX Disclosure: Page 14: 43pp; English.  
 XX the invention relates to a method for inhibiting new tissue growth or  
 CC neointimal formation in blood vessels in a subject that has experienced  
 CC blood vessel injury and preventing exaggerated restenosis in a diabetic  
 CC subject. The method comprises administering an inhibitor of receptor for  
 CC advanced glycation/glycosylation end product (RAGE), so as to inhibit new  
 CC tissue growth or neointimal formation in subject's blood vessels and  
 CC preventing restenosis in the subject. The method is useful for inhibiting  
 CC new tissue growth or neointimal formation in blood vessels in a subject  
 CC like non-human animal, a transgenic non-human animal or a human suffering  
 CC from diabetes, acute thrombotic stroke, venous thrombosis, unstable  
 CC angina, myocardial infarction, abrupt closure following angioplasty or  
 CC stent placement, or thrombosis as a result of peripheral vascular surgery  
 CC the method is also useful for preventing restenosis and for determining  
 CC whether a compound inhibits new tissue growth in a blood vessel in a  
 CC subject. The present sequence is cow receptor for advanced glycosylation  
 CC end product (RAGE) protein.  
 XX SQ Sequence 416 AA;  
 Query Match 86.3%; Score 513.5; DB 23; Length 416;  
 Best Local Similarity 89.1%; Pred. No. 1.5e-46;  
 Matches 98; Conservative 5; Mismatches 6; Indels 1; Caps 1;  
 QY 2 QNITATIGEPVLKCKGAPKPPQPLEWKLNTGRTAEAKVLSPPGSGPWSVAPVLPNGS 61  
 DB 24 QNITATIGEPVLKCKGAPKPPQPLEWKLNTGRTAEAKVLSPPG-GDPWDSVARVLPNGS 82  
 QY 62 LFLPAVGIDEGIERCOAMNNGKTKSNRYRVYVQIPGKPEIVDSASEL 111  
 DB 83 LLLPAVGIDEGIERCOAMNNGKTKSNRYRVYVQIPGKPEIVDPASEL 132  
 RESULT 13  
 AAU77542  
 XX AAU77542 standard; Protein: 416 AA.  
 XX AAU77542;  
 XX 05-JUN-2002 (first entry)  
 XX Bovine receptor for advanced glycosylation end product (RAGE).  
 XX Receptor for advanced glycation end product, RAGE, receptor,  
 KW amyloid beta peptide, blood-brain barrier, neurovascular stress,  
 KW cerebral vasoconstriction, suppression, cerebral blood flow enhanced,  
 KW cerebral amyloid angiopathy; transgenic animal; amyloid beta precursor;  
 KW Alzheimer's disease; Down's syndrome, head trauma, stroke, bovine.  
 XX Hos taurus.  
 XX WO200214519-A1  
 XX 21 FEB 2002  
 XX 14 AUG 2003; 2001WO-US2036.  
 XX 13 OCT 2000; 2000US-0687528.

PR 14-AUG-2000; 2000US-0636648  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX Stera DM, Schmidt AM, Yan SD, Zlokovic B;  
 XX WPI: 2002-257610/30.  
 XX N-PSDB: ABK10855.  
 XX Ameliorating neurovascular stress and decreasing cerebral  
 PT vasoconstriction in subject suffering from chronic/acute cerebral  
 PT amyloid angiopathy, by administering inhibitor of receptor for advanced  
 PT glycation endproduct  
 XX Disclosure: Page 13-14: 68pp; English.  
 XX The invention describes a method of ameliorating neurovascular stress,  
 CC and decreasing cerebral vasoconstriction in subject suffering from  
 CC chronic or acute cerebral amyloid angiopathy, comprising administering  
 CC an inhibitor (I) of receptor for advanced glycation end product (RAGE).  
 CC (I) inhibits transcytosis of amyloid beta peptides across blood-brain  
 CC barrier, thus decreasing cerebral vasoconstriction and increasing  
 CC cerebral blood flow. (I) is useful for treating amyloid angiopathy in a  
 CC subject, decreasing cerebral vasoconstriction in a transgenic non-human  
 CC animal (preferably, transgenic mouse overexpressing mutant human amyloid  
 CC beta precursor protein) or a human, suffering from chronic or acute  
 CC cerebral amyloid angiopathy, preferably, Alzheimer's disease, and for  
 CC ameliorating neurovascular stress comprising cerebral amyloid angiopathy  
 CC in a subject, where the neurovascular stress is caused by Alzheimer's  
 CC disease, aging, Down's syndrome, head trauma or stroke. This is the  
 CC amino acid sequence of bovine receptor for advanced glycation end  
 CC product (RAGE) described in the invention.  
 XX SQ Sequence 416 AA;  
 Query Match 86.3%; Score 513.5; DB 23; Length 416;  
 Best Local Similarity 89.1%; Pred. No. 1.5e-46;  
 Matches 98; Conservative 5; Mismatches 6; Indels 1; Caps 1;  
 QY 2 QNITATIGEPVLKCKGAPKPPQPLEWKLNTGRTAEAKVLSPPGSGPWSVAPVLPNGS 61  
 DB 24 QNITATIGEPVLKCKGAPKPPQPLEWKLNTGRTAEAKVLSPPG-GDPWDSVARVLPNGS 82  
 QY 62 LFLPAVGIDEGIERCOAMNNGKTKSNRYRVYVQIPGKPEIVDSASEL 111  
 DB 83 LLLPAVGIDEGIERCOAMNNGKTKSNRYRVYVQIPGKPEIVDPASEL 132  
 RESULT 14  
 AAU77542  
 XX AAU77542 standard; Protein: 403 AA.  
 XX AAU77542;  
 XX 27-AUG-2002 (first entry)  
 XX Mouse receptor for advanced glycosylation end product (RAGE) protein.  
 XX Mouse: Receptor for advanced glycosylation end product; RAGE; cardiac;  
 KW tissue growth, neointimal formation, blood vessel, restenosis; diabetes,  
 KW myocardial infarction; angioplasty; peripheral vascular surgery; angina;  
 KW transgenic animal, acute thrombotic stroke, venous thrombosis.  
 XX Mus musculus.  
 XX WO200230889 A2.  
 XX 18-APR-2002.  
 XX 12-OCT-2001, 2001WO-US32036.  
 XX 13 OCT 2000; 2000US-0687528.

PA (UYCO) UNIV COLUMBIA NEW YORK.

XX Stern DM, Schmidt AM, Marso S, Topol EJ, Lincott AM.

XX WPT: 2002 4:26:40/45.

XX N PSDB: AAD46954.

XX Inhibiting new tissue growth or neointimal formation in blood vessels

XX of subject suffering from diabetes, stroke and preventing restenosis,

XX comprises administering inhibitor of receptor for advanced glycation end

XX product

XX Disclosure: Page 18; 4 qpp, English.

XX The invention relates to a method for inhibiting new tissue growth or

XX neointimal formation in blood vessels in a subject that has experienced

XX blood vessel injury and preventing exaggerated restenosis in a diabetic

XX subject. The method comprises administering an inhibitor of receptor for

XX advanced glycation/lysosylation end product (RAGE), so as to inhibit new

XX tissue growth or neointimal formation in subject's blood vessels and

XX preventing restenosis in the subject. The method is useful for inhibiting

XX new tissue growth or neointimal formation in blood vessels in a subject

XX like non-human animal, a transgenic non-human animal or a human suffering

XX from diabetes, acute thrombotic stroke, venous thrombosis, unstable

XX angina, myocardial infarction, abrupt closure following angioplasty or

XX stent placement, or thrombosis as a result of peripheral vascular surgery.

XX The method is also useful for preventing restenosis and for determining

XX whether a compound inhibits new tissue growth in a blood vessel in a

XX subject. The present invention is mouse receptor for advanced

XX glycosylation end product (RAGE) protein.

XX Sequence 403 AA:

XX

XX Query Match 86.1% Score 512.5; DB 24; Length 403;

XX Best Local Similarity 87.4%; Pred. No. 1,46 46;

XX Matches 97; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 2 UNITARIGEPVLKCKKAKKKPKQRLKWLKNTGRTFAWKVLSPQGGGWDSVAKVLPNCS 61

DB 24 UNITARIGEPVLKCKKAKKKPKQRLKWLKNTGRTFAWKVLSPQGGGWDSVAKVLPNCS 82

QY 62 LFLPAVGIDGEGTFRKATNKRKEVKSNNYPVPVYQIPKPELVDSASELT 112

DB 83 LLLPATGIVDEGTFRKATNKRKEVKSNNYPVPVYQIPKPELVDSASELT 134

RESULT 15

AAU77544

XX AAU77544 standard; Protein: 403 AA.

XX AAU77544;

XX 05-JUN 2002 (first entry)

XX Marine receptor for advanced glycosylation end product (RAGE).

XX Receptor for advanced glycation end product; RAGE; receptor;

XX amyloid beta peptide; blood brain barrier; neurovascular stress;

XX cerebral vasoconstriction suppressor; cerebral blood flow enhancer;

XX cerebral amyloid angiopathy; transgenic animal; amyloid beta precursor;

XX Alzheimer's disease; Down's syndrome; head trauma; stroke; mouse.

XX Mus musculus.

XX W020024519-A1.

XX 21 FEB 2002.

XX 14 AUG 2001; 2001W00825416.

XX 14-AUG 2000; 200001S 09-4648.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Stern DM, Schmidt AM, Yan SB, Zlokovic B.

XX WPT: 2002 2:36:19/40.

XX N PSDB: ARI0857; ARI0858.

XX Ameliorating neurovascular stress and decreasing cerebral

XX amyloid angiopathy in subject suffering from chronic/acute cerebral

XX amyloid angiopathy, by administering inhibitor of receptor for advanced

XX glycation endproduct

XX Disclosure: Page 17 18; 68pp; English.

XX The invention describes a method of ameliorating neurovascular stress,

XX and decreasing cerebral vasoconstriction in subject suffering from

XX chronic or acute cerebral amyloid angiopathy, comprising administering

XX an inhibitor (I) of receptor for advanced glycation end product (RAGE).

XX (I) inhibits transcytosis of amyloid beta peptides across blood brain

XX barrier, thus decreasing cerebral vasoconstriction and increasing

XX cerebral blood flow. (I) is useful for treating amyloid angiopathy in a

XX subject, decreasing cerebral vasoconstriction in a transgenic non human

XX animal (preferably, transgenic mouse overexpressing mutant human amyloid

XX beta precursor protein) or a human, suffering from chronic or acute

XX cerebral amyloid angiopathy, preferably, Alzheimer's disease, and for

XX ameliorating neurovascular stress in subject suffering from amyloid angiopathy

XX in a subject, where the neurovascular stress is caused by Alzheimer's

XX disease, aging, Down's syndrome, head trauma or stroke. This is the

XX amino acid sequence of murine receptor for advanced glycation end

XX product (RAGE) described in the invention.

XX Sequence 403 AA:

XX

XX Query Match 86.1% Score 512.5; DB 24; Length 403;

XX Best Local Similarity 87.4%; Pred. No. 1,46 46;

XX Matches 97; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 2 UNITARIGEPVLKCKKAKKKPKQRLKWLKNTGRTFAWKVLSPQGGGWDSVAKVLPNCS 61

DB 24 UNITARIGEPVLKCKKAKKKPKQRLKWLKNTGRTFAWKVLSPQGGGWDSVAKVLPNCS 82

QY 62 LFLPAVGIDGEGTFRKATNKRKEVKSNNYPVPVYQIPKPELVDSASELT 112

DB 83 LLLPATGIVDEGTFRKATNKRKEVKSNNYPVPVYQIPKPELVDSASELT 134

Search Completed: May 30, 2003, 15:55:04

Job Time : 27.4775 secs

GenCore version 5.1.1 b  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein protein search, using sw model

Run on: May 30, 2003, 15:53:32 ; Search time 8.82883 seconds  
(without alignments)  
1284.095 Million cell updates/sec

Title: US-09-872-185B-1

Perfect score: 595

Sequence: 1 AONITARIKGEPLVLCCKGAP...VRVYQIIPCKPIVDASSELT 112

Scoring table: BLOSUM62

Gapop 10 0, Gapex 6.5

Seatched: 383519 seqs, 101223694 residues

Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pcp.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/FCI\_NEW\_PUB.pcp.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pcp.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pcp.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pcp.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pcp.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pcp.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pcp.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pcp.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	595	100.0	112	10	US-09-872-185B-1
2	595	100.0	332	9	US-09-851-071-1
3	595	100.0	332	10	US-09-872-185B-2
4	591	99.3	405	1	US-08-755-235-2
5	513	86.3	416	1	US-08-755-235-2
6	162	27.2	30	8	US-08-948-131-1
7	162	27.2	30	9	US-08-948-131-1
8	162	27.2	30	10	US-09-851-071-5
9	162	27.2	30	10	US-09-872-185B-3
10	149	25.0	30	8	US-08-948-131-2
11	149	25.0	30	10	US-09-872-185B-7
12	145	24.4	30	8	US-08-948-131-4
13	145	24.4	30	10	US-09-872-185B-4
14	133	22.4	30	8	US-08-948-131-3
15	133	22.4	30	10	US-09-872-185B-6
16	104	17.5	2586	10	US-09-872-185B-5
17	104	17.5	2586	10	US-09-905-129-11
18	104	17.5	2586	10	US-09-905-129-14
19	104	17.5	2586	10	US-09-991-630-11
20	104	17.5	2586	10	US-09-991-630-14

ALIGNMENTS

RESULT 1

US-09-872-185B-1  
; Sequence 1, Application US/09872185B  
; Patent No. US2002017550A1  
; GENERAL INFORMATION:  
; APPLICANT: Stern, David M.  
; APPLICANT: Herold, Kevin  
; APPLICANT: Yan, Shi Du  
; APPLICANT: Schmidt, Ann Marie  
; APPLICANT: Lamster, Ira  
; TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION  
; FILER REFERENCE: 0575/64080  
; CURRENT APPLICATION NUMBER: US/09/872,185B  
; CURRENT FILING DATE: 2001-06-01  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 112  
; TYPE: PRT  
; ORGANISM: Human  
US-09-872-185B-1

Query Match	100.0%	Score	595	DR	10	Length	112
Rest Local Similarity	100.0%	Pred. No	1	le-49			
Matches 112	Conservative	0	Mismatches	0	Indels	0	Gaps
0							

Qy 1 AONTARIGEPVLKCKGAEKPPQRLWKINTGRTEAKVKLSPGGGPDVSVARVLPNG 60  
DB 1 AONTARIGEPVLKCKGAEKPPQRLWKINTGRTEAKVKLSPGGGPDVSVARVLPNG 60  
Qy 61 SLFLPAPVGIODGIFHCQAMNPNKRTKSNRYRVVYQIIPCKPIVDASSELT 112  
Db 61 SLFLPAPVGIODGIFHCQAMNPNKRTKSNRYRVVYQIIPCKPIVDASSELT 112

RESULT 2

US-09-851-071-1  
; Sequence 1, Application US/09851071  
; Patent No. US2002017550A1  
; GENERAL INFORMATION:  
; APPLICANT: Schmidt, Anne Marie  
; APPLICANT: Stern, David

```

? TITLE OF INVENTION: A METHOD FOR INHERITING TUMOR INVASION OR SPREADING IN A SUBJECT
? FILE NUMBER: 05/09/872-185b-1.rapb
? CURRENT APPLICATION NUMBER: US/09/872-185b-1
? CURRENT FILING DATE: 2001 05 08
? NUMBER OF SEQ ID NOS: 6
? SOFTWARE: Patent in version 4.1
? SEQ ID NO 1
? LENGTH: 432
? TYPE: PRT
? ORGANISM: Human
US-09-872-185b-1

```

```

Query Match
Best Local Similarity 100.0% Score 595; DB 1; Length 432;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTATIGEPVLKCGAKPKPQRIEWNKUNTEAKVLSPPGGGPDWSVARVLPNG 60
DB 1 AONTATIGEPVLKCGAKPKPQRIEWNKUNTEAKVLSPPGGGPDWSVARVLPNG 60
QY 61 SLFLPAVGIGDEGIFRCQAMNNGKETSNNYRVVYQIPKPEIVDSASELT 112
DB 61 SLFLPAVGIGDEGIFRCQAMNNGKETSNNYRVVYQIPKPEIVDSASELT 112

```

```

RESULT 4
US-09-872-185b-2
? Sequence 2, April 11, 2001 05/09/872-185b
? Patent No. US2000059423A1
? GENERAL INFORMATION:
? APPLICANT: Stern, David M.
? APPLICANT: Herold, Kevin
? APPLICANT: Yan, Shi Du
? APPLICANT: Schmidt, Ann Marie
? APPLICANT: Lamster, Ira
? TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
? FILE REFERENCE: 0575/50159
? CURRENT APPLICATION NUMBER: US/09/872-185b
? CURRENT FILING DATE: 2001 06 01
? NUMBER OF SEQ ID NOS: 16
? SOFTWARE: Patent in version 4.1
? SEQ ID NO 2
? LENGTH: 432
? TYPE: PRT
? ORGANISM: Human
US-09-872-185b-2

```

```

Query Match
Best Local Similarity 100.0% Score 595; DB 1; Length 432;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTATIGEPVLKCGAKPKPQRIEWNKUNTEAKVLSPPGGGPDWSVARVLPNG 60
DB 1 AONTATIGEPVLKCGAKPKPQRIEWNKUNTEAKVLSPPGGGPDWSVARVLPNG 60
QY 61 SLFLPAVGIGDEGIFRCQAMNNGKETSNNYRVVYQIPKPEIVDSASELT 112
DB 61 SLFLPAVGIGDEGIFRCQAMNNGKETSNNYRVVYQIPKPEIVDSASELT 112

```

```

RESULT 4
US-09-872-185b-4
? Sequence 4, Application US/09/872-185b
? Patent No. US2000059423A1
? GENERAL INFORMATION:
? APPLICANT: Stern, David M.
? APPLICANT: Schmidt, Ann Marie
? APPLICANT: Yan, Shi Du
? TITLE OF INVENTION: METHODS FOR TREATING SYMPTOMS OF DIABETES
? FILE REFERENCE: 0575/50159
? CURRENT APPLICATION NUMBER: US/09/872-185b
? CURRENT FILING DATE: 1996 11 22
? NUMBER OF SEQ ID NOS: 4

```

```

? SOFTWARE: Patent in version 4.1
? SEQ ID NO 4
? LENGTH: 405
? TYPE: PRT
? ORGANISM: Human
US-08-755-235-4

```

```

Query Match
Best Local Similarity 99.1% Score 591; DB 1; Length 405;
Matches 111; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 AONTATIGEPVLKCGAKPKPQRIEWNKUNTEAKVLSPPGGGPDWSVARVLPNG 60
DB 23 AONTATIGEPVLKCGAKPKPQRIEWNKUNTEAKVLSPPGGGPDWSVARVLPNG 82
QY 61 SLFLPAVGIGDEGIFRCQAMNNGKETSNNYRVVYQIPKPEIVDSASELT 112
DB 83 SLFLPAVGIGDEGIFRCQAMNNGKETSNNYRVVYQIPKPEIVDSASELT 144

```

```

RESULT 5
US-08-755-235-2
? Sequence 2, April 11, 2001 05/09/872-185b
? Patent No. US2000059423A1
? GENERAL INFORMATION:
? APPLICANT: Stern, David M.
? APPLICANT: Schmidt, Ann Marie
? APPLICANT: Wu, Jun
? TITLE OF INVENTION: METHOD FOR TREATING SYMPTOMS OF DIABETES
? FILE REFERENCE: 0575/50159
? CURRENT APPLICATION NUMBER: US/08/755-235
? CURRENT FILING DATE: 1996-11-22
? NUMBER OF SEQ ID NOS: 4
? SOFTWARE: Patent in version 4.1
? SEQ ID NO 2
? LENGTH: 416
? TYPE: PRT
? ORGANISM: Bovine
US-08-755-235-2

```

```

Query Match
Best Local Similarity 86.4% Score 513.5; DB 1; Length 416;
Matches 98; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 QNTATIGEPVLKCGAKPKPQRIEWNKUNTEAKVLSPPGGGPDWSVARVLPNG 61
DB 24 QNTATIGEPVLKCGAKPKPQRIEWNKUNTEAKVLSPPGGGPDWSVARVLPNG 82
QY 62 SLFLPAVGIGDEGIFRCQAMNNGKETSNNYRVVYQIPKPEIVDSASELT 111
DB 83 SLFLPAVGIGDEGIFRCQAMNNGKETSNNYRVVYQIPKPEIVDSASELT 142

```

```

RESULT 5
US-08-948-131-1
? Sequence 1, Application US/08/948-131
? Patent No. US20010954357A1
? GENERAL INFORMATION:
? APPLICANT: Stern, David
? APPLICANT: Yan, Shi Du
? APPLICANT: Schmidt, Ann Marie
? TITLE OF INVENTION: Method Hinding Site of Kato and Uses
? NUMBER OF SEQUENCES: 5
? CORRESPONDENCE ADDRESS:
? ADDRESS: Cooper & Dunham
? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10046
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk

```

1. CURRENT APPLICATION NUMBER

```

1 APPLICATION NUMBER: US/08/9448,141
2
3 FILING DATE: 09-OCT 1997
4
5 CLASSIFICATION: 530
6
7 ATTORNEY/AGENT INFORMATION:
8
9 NAME: White, John P
10
11 REGISTRATION NUMBER: 28,678
12
13 REFERENCE/DOCKET NUMBER: 53447
14
15 TELECOMMUNICATION INFORMATION:
16
17 TELEPHONE: 212-278-0400
18
19 TELEFAX: 212-391-0526
20
21 INFORMATION FOR SEQ ID NO: 2:
22
23 SEQUENCE CHARACTERISTICS:
24
25 LENGTH: 40 amino acids
26
27 TYPE: amino acid
28
29 STRANDNESS: single
30
31 TOPOLOGY: linear
32
33 MOLECULE TYPE: peptide
34
35 DS: 08-9448-141.2

```

Query Match	25.0%	Score 149;	DB 8;	Length 30;
Best Local Similarity	94.1%	Prod. No. 8 to 08;		
Matches	27;	Conservative	1;	Indels 0;
		Mismatches	1;	Gaps 0;

**Qy**      2 QNITARIGEPLVLCKCGAPKKPPQRLEWK 30  
          | | | | | | | | | | | | : | | |  
**Dd**      2 QNITARIGEPLVLS\*KGAPKKPPQQLEWK 30

```

> RESULT 11
> US-09-872 185B 4
> Sequence 4, Application US/09872185B
> Patent No. US2020122799A1
> GENERAL INFORMATION:
> APPLICANT: Stern, David M.
> APPLICANT: Hetold, Kevan
> APPLICANT: Yan, Shi Du
> APPLICANT: Schmidt, Ann Marie
> APPLICANT: Lanster, Ira
> TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
> FILE REFERENCE: 0575/64080
> CURRENT APPLICATION NUMBER: US/09/872,185B
> CURRENT FILING DATE: 2001-06-01
> NUMBER OF SEQ ID NOS: 16
> SOFTWARE: PatentIn version 3.1
> SEQ ID NO 4
> LENGTH: 30
> TYPE: PRT
> ORGANISM: Murine
> US-09-872-185B-4

```

Query Match	25.08	score 149;	DB 10;	length 40;
Best Local Similarity	93.18	pred. NO. 8.3e-08;		
Matches	27	Conservative	1	Mismatches 1;
				Indels 0;
				Gaps 0;

[illegible]

```

1 RESULT 12
2 US-08-948-131-4
3 ? Sequence 4, Application US/08948131
4 ? Patent No. US20010054357A1
5 ? GENERAL INFORMATION:
6 ? APPLICANT: Stern, David
7 ? APPLICANT: Yan, Shi Du
8 ? APPLICANT: Schmidt, Ann Marie
9 ? TITLE OF INVENTION: Liquid Binding Site of Kape and Uses
10 ?
11 ? TITLE OF INVENTION: Hicroot
12 ?
13 ? NUMBER OF SEQUENCES: 5
14 ?
15 ? CORRESPONDENCE ADDRESS:
16 ?
17 ? ADDRESSEE: Cooper & Dunham
18 ?
19 ? STREET: 1185 Avenue of the Americas

```

CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.0.40  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/94B,131  
FILING DATE: 09-OCT 1997  
CLASSIFICATION: 5,40  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 5,447  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212 391 0526  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 30 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
P.S. 08-94B-131-4

Query Match 24.4%; Score 145; bits 8; Length 40;  
Best Local Similarity 89.7%; Pred. No. 2e 07;  
Matches 26; Conservative 2; Mismatches 1, indels 0; gaps 0;

Qy 2 QNITARIGEPILVI,KCKGAI-KKPPQRLWK 30  
 1111111:111 1111111:111

Dp 2 QNITARIGKPLVINCCKGAEKKPPQLEWK 30  
 1111111:111 1111111:111

```

1  RESULT 13
2  US-09-872 185H 6
3  ? Sequence 6, Application US/09872185H
4  ? Patent No. US2002012799A1
5  ? GENERAL INFORMATION:
6  ? APPLICANT: Stern, David M.
7  ? APPLICANT: Herold, Kevin
8  ? APPLICANT: Yan, Shi Da
9  ? APPLICANT: Schmidt, Ann Marie
10 ? TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
11 ? FILE REFERENCE: 0575/64080
12 ? CURRENT APPLICATION NUMBER: US/09/872,185H
13 ? CURRENT FILING DATE: 2001-06 01
14 ? NUMBER OF SEQ ID NOS: 16
15 ? SOFTWARE: PatentIn version 3.1
16 ? SEQ ID NO 6
17 ? LENGTH: 36
18 ? TYPE: PRT
19 ? ORGANISM: bovine
20 ? US-09-872-185H 6

```

Query Match	24.4%	Score 145;	108 10;	Length 40;
Best Local Similarity	89.7%	Pred. No. 20 07;		
Matches	26, Conservative	2; Mismatches	1;	Indices 0; Gaps 0;

QY	2	UNITARI	CEPL	VL	NK	GAP	PP	Q	LEW	30
QY	2	UNITARI	CEPL <td>VL <td>NK <td>GAP <td>PP <td>Q <td>LEW <td>30</td> </td></td></td></td></td></td>	VL <td>NK <td>GAP <td>PP <td>Q <td>LEW <td>30</td> </td></td></td></td></td>	NK <td>GAP <td>PP <td>Q <td>LEW <td>30</td> </td></td></td></td>	GAP <td>PP <td>Q <td>LEW <td>30</td> </td></td></td>	PP <td>Q <td>LEW <td>30</td> </td></td>	Q <td>LEW <td>30</td> </td>	LEW <td>30</td>	30
QY	2	UNITARI	CEPL <td>VL <td>NK <td>GAP <td>PP <td>Q <td>LEW <td>30</td> </td></td></td></td></td></td>	VL <td>NK <td>GAP <td>PP <td>Q <td>LEW <td>30</td> </td></td></td></td></td>	NK <td>GAP <td>PP <td>Q <td>LEW <td>30</td> </td></td></td></td>	GAP <td>PP <td>Q <td>LEW <td>30</td> </td></td></td>	PP <td>Q <td>LEW <td>30</td> </td></td>	Q <td>LEW <td>30</td> </td>	LEW <td>30</td>	30
QY	2	UNITARI	CEPL <td>VL <td>NK <td>GAP <td>PP <td>Q <td>LEW <td>30</td> </td></td></td></td></td></td>	VL <td>NK <td>GAP <td>PP <td>Q <td>LEW <td>30</td> </td></td></td></td></td>	NK <td>GAP <td>PP <td>Q <td>LEW <td>30</td> </td></td></td></td>	GAP <td>PP <td>Q <td>LEW <td>30</td> </td></td></td>	PP <td>Q <td>LEW <td>30</td> </td></td>	Q <td>LEW <td>30</td> </td>	LEW <td>30</td>	30

RESULT 14  
US-08-948 141-3  
; Sequence 3, Application US/08948141  
; Patent No. US20010053357A1



GENERAL INFORMATION:  
 APPLICANT: Stern, David  
 APPLICANT: Yan, Shi Du  
 APPLICANT: Schmidt, Ann Marie  
 TITLE OF INVENTION: Ligand Binding Site of Rage and Uses  
 TITLE OF INVENTION: Thersol  
 NUMBER OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Cooper & Dunham  
 STREET: 1185 Avenue of the Americas  
 CITY: New York  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10036  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/948,131  
 FILING DATE: 09-OCT-1997  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: White, John P  
 REGISTRATION NUMBER: 26,678  
 REFERENCE/DOCKET NUMBER: 53447  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-278-0400  
 TELEFAX: 212-391-0536  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 30 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-948-131-3

Query Match: 22.4%, Score 133, Pos 8, Length 30;  
 Best Local Similarity: 82.8%, Pred. No. 2.8e-06;  
 Matches: 24; Conservative: 2; Mismatches: 3; Indels: 0; Gaps: 0;

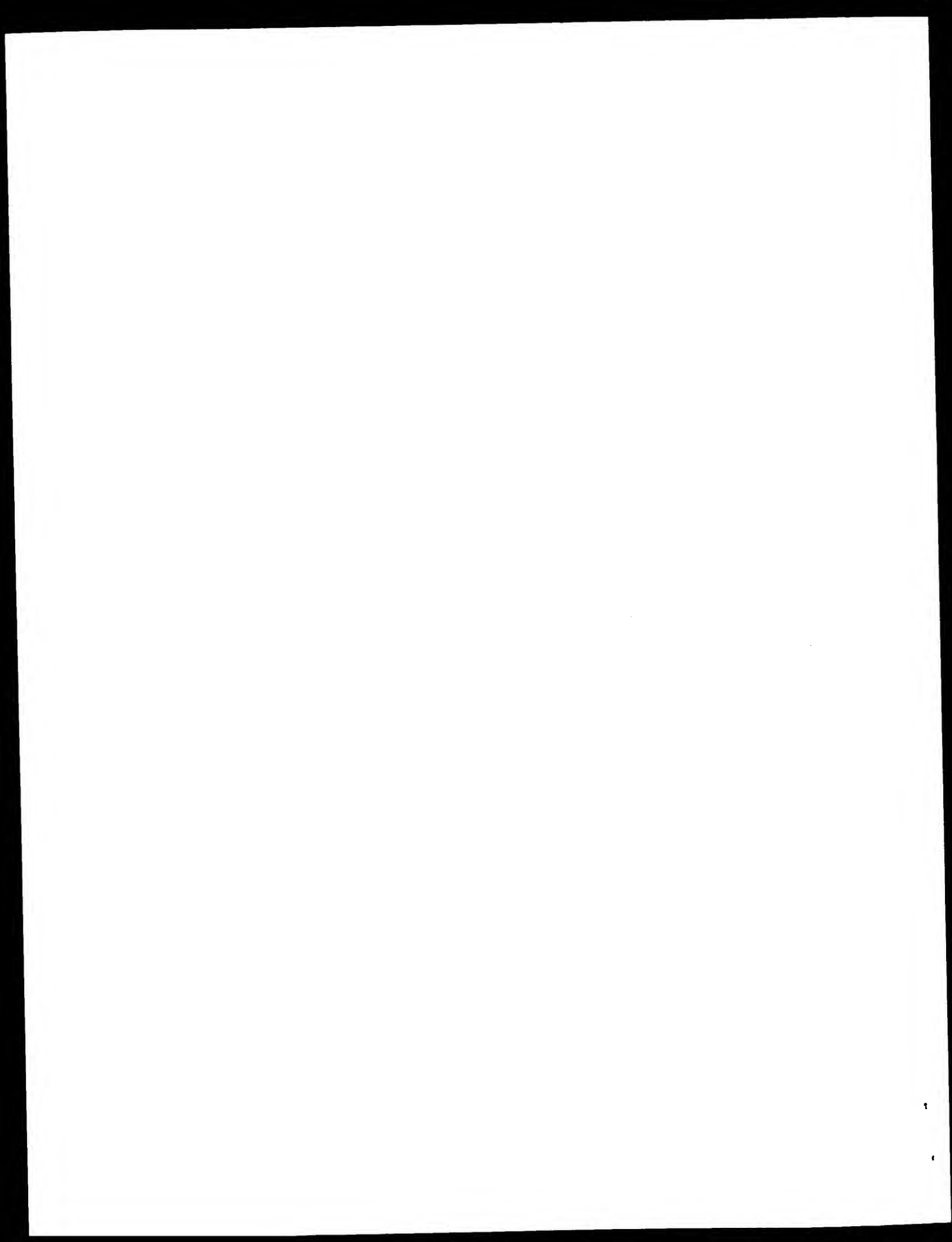
QY 2 QNITARIGEPLVKCKGAPKPPQRLWK 30  
 DB 2 QNITARIGEPLVKCKGAPKPPQRLWK 30

RESULT 15  
 US-09-872-185B-5  
 Sequence 5, Application US/09872185B  
 Patent No. US20020122799A1  
 GENERAL INFORMATION:  
 APPLICANT: Stern, David M.  
 APPLICANT: Herold, Kevin  
 APPLICANT: Yan, Shi Du  
 APPLICANT: Schmidt, Ann Marie  
 APPLICANT: Lamster, Ira  
 TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION  
 FILE REFERENCE: 0575/64080  
 CURRENT APPLICATION NUMBER: US/09/872,185B  
 CURRENT FILING DATE: 2001-06-01  
 NUMBER OF SEQ ID NOS: 16  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 5  
 LENGTH: 30  
 TYPE: PRT  
 ORGANISM: Rat  
 US-09-872-185B-5

Query Match: 22.4%, Score 133, Pos 10, Length 30;  
 Best Local Similarity: 82.8%, Pred. No. 2.8e-06;  
 Matches: 24; Conservative: 2; Mismatches: 3; Indels: 0; Gaps: 0;

QY 2 QNITARIGEPLVKCKGAPKPPQRLWK 30  
 DB 2 QNITARIGEPLVKCKGAPKPPQRLWK 30

Search completed: May 30, 2003, 15:59:06  
 Job time: 9.92833 secs



Result No.	Query Match	Score	Length	DB	ID	Description	
1	595	100.0	404	1	T01596	advanced glycosyltransferase	
2	516.5	86.8	492	2	T09062	probable advanced glycosyltransferase	
3	513.5	86.3	416	1	A42879	advanced glycosyltransferase	
4	114	19.2	1447	2	A54100	tumor suppressor protein	
5	108	18.2	267	2	A94482	probable tumor suppressor protein	
6	103	17.3	1147	2	A59407	myosin-light-chain	
7	101	17.0	1176	2	T00583	myosin-light-chain	
8	97	16.3	1040	2	A49356	transient axonal growth cone	
9	96.5	16.2	1257	1	A41060	neural cell adhesion molecule	
10	96	16.1	1018	2	J04211	neural cell adhesion molecule	
11	96	16.1	1040	2	A34695	axonal glycoprotein	
12	96	16.1	1144	2	T14316	rig 1 protein - mouse	
13	95.5	16.1	2295	2	G68369	protein unc-52-like	
14	95.5	16.1	3375	2	T19821	hypothetical protein	
15	95	16.0	1892	2	A4667	tumor suppressor protein	
16	95	16.0	1427	2	T1664	tumor suppressor protein	
17	93	15.6	1018	2	A54744	contactin 1 precursor	
18	92.5	15.5	1612	2	T08952	duffy protein - mouse	
19	92.5	15.5	1651	2	T14160	transmembrane receptor	
20	92	15.5	792	2	A63319	calcium-binding protein	
21	91	15.3	523	2	G50478	neuronal cell adhesion molecule	
22	91	15.3	719	2	T09676	neuronal cell adhesion molecule	
23	91	15.3	1535	2	S45224	peroxidasin - fruit fly	
24	89	15.0	1320	2	S05914	neuronal cell surface protein	
25	88	14.8	1332	2	T06067	pregnancy-specific cell adhesion protein	
26	88	14.9	1013	2	S10247	pregnancy-specific cell adhesion protein	
27	87	14.6	210	2	J14152	pregnancy-specific cell adhesion protein	
28	87	14.6	273	2	G28938	pregnancy-specific cell adhesion protein	
29	87	14.6	275	2	A28938	pregnancy-specific cell adhesion protein	

A:Residues: 23-24, 'X', 26, 37, 'X', 34-49, 'XX', 52, 54, 'SCB'.  
 A:Experimental source: endothelial cells.  
 A:Note: Sequence extracted from NCHI backbone (NCBIP:109464).  
 C:Comment: Advanced glycosylation end products are heterogeneous nonenzymatically gly-  
 cellular function, thus contributing to tissue lesions in diabetes.  
 C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide  
 ates in the neurotoxic pathway that produces dementia in Alzheimer's disease.  
 C:Function:  
 C:Description: neuronal receptor for amphoterin, a DNA binding protein involved in ne  
 C:Superfamily: advanced glycosylation end products receptor; immunoglobulin homology  
 C:Keywords: Alzheimer's disease; glycoprotein; receptor; transmembrane protein  
 F:1-23/Domain: signal sequence #status predicted -SIG>  
 F:23-416/Product: advanced glycosylation end products receptor RAGE #status predicted  
 F:23-354/Domain: extracellular #status predicted -EXT>  
 F:31-100/Domain: immunoglobulin homology <IM1>  
 F:116-269/Domain: immunoglobulin homology <IM2>  
 F:262-313/Domain: immunoglobulin homology <IM3>  
 F:355-372/Domain: transmembrane #status predicted -TM>  
 F:373-416/Domain: intracellular #status predicted -INT>  
 F:425-80/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:38-98,143-207,269-411/Dissulfide bonds: #status predicted  
 Query Match 86.3%; Score 513.5; DB 1; Length 416;  
 Best Local Similarity 89.1%; Pred. No. 9e 44;  
 Matches 98; Conservative 5; Mismatches 6; Indels 1; Gaps 1;  
 QY 2 QNITARIGEPVLVKCKGAPKKPPQRLKWLKNTGRTEAKVKVLSPOGGGHWDSVARVLPGNS 61  
 DB 24 QNITARIGEPVLVKCKGAPKKPPQRLKWLKNTGRTEAKVKVLSPOGGGHWDSVARVLPGNS 82  
 QY 62 LFLPAVGIDQEGIFRCQAMNRNGKTKSNRYRVRVYQIPCKPEIVDSASELT 112  
 DB 83 LFLPAVGIDQEGIFRCQAMNRNGKTKSNRYRVRVYQIPCKPEIVDSASELT 142  
 RESULT 4  
 A54100  
 tumor suppressor protein DCC precursor human  
 N:Alternate names: colorectal cancer suppressor DCC  
 C:Species: Homo sapiens (man)  
 C:Date: 02-Aug-1994 #sequence\_revision 02-Aug-1994 #text\_change 05-Nov-1999  
 C:Accession: A54100; A40098  
 R:Bedrick, L.; Cho, K.F.; Fearon, F.R.; Wu, T.C.; Kinzler, K.W.; Vogelstein, B.  
 Genes Dev. 8, 1174-1183, 1994  
 A:Title: The DCC gene product in cellular differentiation and colorectal tumorigenesis  
 A:Reference number: A54100; MUID:95011532; PMID:7926722  
 A:Accession: A54100  
 A:Molecule type: mRNA  
 A:Residues: 1-1447 <DED>  
 A:Cross-references: EMBL:X76132; NID:q454209; PIDN:CAA43745.1; PID:q454210  
 R:Pearson, E.R.; Cho, K.R.; Nigro, J.M.; Kern, S.F.; Simons, J.W.; Huppert, J.M.; Hami  
 Science 247, 49-56, 1990  
 A:Title: Identification of a chromosome 18q gene that is altered in colorectal cancer  
 A:Reference number: A40098; MUID:90100559; PMID:2294591  
 A:Accession: A40098  
 A:Molecule type: mRNA  
 A:Residues: 1-750 <PFA>  
 A:Cross-references: GB:M32292; NID:g181492; PIDN:AAA45751.1; PID:g181494  
 C:Genetics:  
 A:Gene: GDB:DCC  
 A:Cross-references: GDB:119838; OMIM:120470  
 A:Map position: 18q21.1-18q21.1  
 C:Keywords: transmembrane protein; tumor suppressor  
 F:1-25/Domain: signal sequence #status predicted -SIG>  
 F:26-1447/Product: tumor suppressor protein DCC #status predicted -MAT>  
 Query Match 19.2%; Score 114; DB 2; Length 1447;  
 Best Local Similarity 31.3%; Pred. No. 0.0047;  
 Matches 31; Conservative 20; Mismatches 48; Indels 10; Gaps 4;  
 QY 2 QNITARIGEPVLVKCKGAPKKPPQRLKWLKNTGRTEAKVKVLSPOGGGHWDSVARVLPGNS 61  
 DB 147 ESVTAFMGDTVLKCE-VIGEPMPTHWQKNO -QULTF1 -FCDSKVVVLPSSGA 196

Best Local Similarity 100.0%; Pred. No. 6e-52;  
 Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AONTARIGEPVLVKCKGAPKKPPQRLKWLKNTGRTEAKVKVLSPOGGGHWDSVARVLPGNS 60  
 DB 23 AONTARIGEPVLVKCKGAPKKPPQRLKWLKNTGRTEAKVKVLSPOGGGHWDSVARVLPGNS 82  
 QY 61 SLELPAVGIDQEGIFRCQAMNRNGKTKSNRYRVRVYQIPCKPEIVDSASELT 112  
 DB 83 SLELPAVGIDQEGIFRCQAMNRNGKTKSNRYRVRVYQIPCKPEIVDSASELT 134  
 RESULT 2  
 T09062  
 probable advanced glycosylation end-products receptor precursor - mouse  
 N:Alternate names: RAGE  
 C:Species: Mus musculus (house mouse)  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jan-2000  
 C:Accession: T09062  
 R:Rowen, L.; Mahairas, G.; Qin, S.; Ahern, M.E.; Dankers, C.; Lasky, S.; Loretz, C.; So  
 submitted to the EMBL data library, October 1997  
 A:Description: Sequence of the mouse major histocompatibility locus class III region.  
 A:Reference number: Z16543  
 A:Accession: T09062  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-402 <ROW>  
 A:Cross-references: EMBL:AF030001; NID:g2564945; PID:g2564950  
 C:Genetics:  
 A:Gene: RAGE  
 A:Map position: 17  
 A:Insertions: 187; 534; 118; 1, 2, 3, 4, 169; 1, 209; 1, 272; 3, 320; 1, 471; 2  
 C:Superfamily: advanced glycosylation end products receptor; immunoglobulin homology  
 C:Keywords: receptor; transmembrane protein  
 F:31-100/Domain: immunoglobulin homology <IM>  
 Query Match 86.8%; Score 516.5; DB 2; Length 402;  
 Best Local Similarity 88.3%; Pred. No. 4.3e-44;  
 Matches 98; Conservative 3; Mismatches 9; Indels 1; Gaps 1;  
 QY 2 QNITARIGEPVLVKCKGAPKKPPQRLKWLKNTGRTEAKVKVLSPOGGGHWDSVARVLPGNS 61  
 DB 24 QNITARIGEPVLVKCKGAPKKPPQRLKWLKNTGRTEAKVKVLSPOGGGHWDSVARVLPGNS 82  
 QY 62 LFLPAVGIDQEGIFRCQAMNRNGKTKSNRYRVRVYQIPCKPEIVDSASELT 112  
 DB 83 LFLPAVGIDQEGIFRCQAMNRNGKTKSNRYRVRVYQIPCKPEIVDSASELT 133  
 RESULT 3  
 A42879  
 advanced glycosylation end-products receptor precursor - bovine  
 N:Alternate names: advanced glycosylation end product-binding protein, 35K; glycoprotein  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 04-Mar-1993 #sequence\_revision 07-Feb-1997 #text\_change 16-Jul-1999  
 C:Accession: A42879; A42878; S27949  
 R:Neuper, M.; Schmidt, A.M.; Bretz, J.; Yan, S.D.; Wang, F.; Pan, Y.C.; Elliston, K.; Su  
 J. Biol. Chem. 267, 14998-15004, 1992  
 A:Title: Cloning and expression of a cell surface receptor for advanced glycosylation en  
 A:Reference number: A42879; MUID:92340547; PMID:1378843  
 A:Accession: A42879  
 A:Molecule type: mRNA  
 A:Residues: 1-416 <NEE>  
 A:Cross-references: GB:M91212; NID:g163650; PIDN:AAA03575.1; PID:g163651  
 A:Experimental source: lung  
 A:Note: sequence extracted from NCHI backbone (NCBIP:109466)  
 A:Note: parts of this sequence, including the amino end of the mature protein, were dete  
 R:Schmidt, A.M.; Vianna, M.; Gerlach, M.; Bretz, J.; Ryan, J.; Kao, J.; Esposito, C.; He  
 J. Biol. Chem. 267, 14987-14997, 1992  
 A:Title: Isolation and characterization of two binding proteins for advanced glycosylat  
 A:Reference number: A42878; MUID:92340546; PMID:1121822  
 A:Accession: A42878  
 A:Molecule type: protein





```

9 GRP.VLKCKGAKKKPQRIEWKLNIGRTIAWKVLSFGGCPWDSVAR-----VLPNG 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
56 GEPATLPCR-AECPQPPNTEWKNGAPV-----ATAPEDPAHPILLPSG 99
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
41 STPLPAVAV-----TLQEGTFCQAMNKNGKETKSNYKRVK-----YUJPG 100
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
100 ALFVPEK.VHGRKSKHDFGVYTCVAKNYTAAASKNASLFAVAVLFDROSPG 151
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13

CBB369

Protein: unc52 (imported) - Caenorhabditis elegans

Species: Caenorhabditis elegans

C.Dat: 4 - 10-May-2001 #suppl\_1 revising 10 May 2001 #next\_change 18-Jul-2001

C.Accession: CBB369; E88369

Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A.Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio

A.Reference: number: P75500; URL: 99269613; PMID:9451916

A.Note: see websites genome.wi.mit.edu/spec/C\_elegans/ and www.sanger.ac.uk/projects/

A:Accession: C88369  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-2295 SGTG  
 A:Cross references: GB chr\_11; PID:G4874868; GSPDB:GN00020  
 A:Accession: E8854  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-2295 SGTG  
 A:Cross references: GB chr\_11; PID:CA8077021; P1: g4861368; GS116:CH00020  
 C:Genetics:  
 A:Gene: unc-52  
 A:Map position: 2  
  
 Query Match  
 Best Local Similarity 25.4%, Pred. No. 0.44,  
 Matches 31; Conservative 21; Mismatches 41; Indels 29; Gaps 5;  
  
 8 IGEPVVKCK--CAKKEDETFEWTINTTETAKVLSDEES--DQTSVAPV DNGST FT 6A

DB	1803	VGRPIQVCKEAFGAPZDPEFEVWJHDPG-----PERGILJDDFKPVTISEQFIRH	1853
QY	65	PAWGLDEGIFRCQAZNNRNGKETKSNY-----RVRYQIRG-----KPEIVDS	107
EL	1854	PNWGLNAGVYCKGSSAHATAIKNIYIEWVEFSKIAIVSLGGSSQWFDQGEKGLICT	1913
QY	108	AS	109
DB	1914	AT	1915

T19821  
hypothetical protein ZC101.2e - *Caenorhabditis elegans*  
C1Species: *Caenorhabditis elegans*  
C1Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 01-Dec-2000  
C1Accession: T19821; T19820; T27490; T27488; T27489; T27487; A47648; B47648;  
K1Bayes, C.  
submitted to the EMBL Data Library, March 1997  
A1Reference number: Z13782  
A1Accession: T19821  
A1Status: preliminary, translated from GB/EMBL/DDBJ  
A1Molecule type: DNA  
A1Residues: 1-3375 <W 1>  
A1Cross references: EMBL:Z1375, rDNA:CA07569.1, GSDB:G000020, CESP:ZC101.2e  
A1Experimental source: clone C3806  
A1Accession: T19819

A:Status: preliminary; translated from GB/EMBL/DDDBJ  
A:Molecule type: DNA  
A:Accession: 1-244; 'E', 'REFL', '3369', 'GN', '3372-3373', 'G', '3375', 'PLRHRRRAGNGDILSRKTRTTTKLP  
A:Cross-references: EMBL:Z93345; 1:DN CARB567.1; QSDH; GN:0020; CESP:ZC101.2a





GenCore version 5.1  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 15:37:16, Search time 6.55856 seconds  
(without alignments)  
708.288 million cell updates/sec

Title: US-09-872-185B-1  
Perfect score: 595  
Sequence: 1 AQNITARIQGEPLVLCCKGAP VFVYQIPKPEIVDSASELT 112

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	595	100.0	404	1	RAGE_HUMAN	Q15109 homo sapien
2	513.5	86.3	416	1	RAGE_BOVIN	Q28173 bos taurus
3	512.5	86.1	403	1	RAGE_MOUSE	Q62151 mus musculu
4	494.5	83.1	402	1	RAGE_RAT	Q63495 rattus norv
5	115	19.3	1447	1	DCC_MOUSE	P70211 mus musculu
6	114	19.2	1447	1	DCC_HUMAN	P43146 homo sapien
7	103	17.3	1147	1	KMIS_PARIT	P29234 oryctolagus
8	101	17.0	1176	1	KMIS_BOVIN	Q28824 bos taurus
9	97	16.3	1040	1	AXOL_HUMAN	Q02246 homo sapien
10	96.5	16.2	1257	1	CAML_HUMAN	P32004 homo sapien
11	96	16.1	1040	1	AXOL_RAT	P22063 rattus norv
12	95.5	16.1	2481	1	UN52_CAEFL	Q06501 caenorhabdi
13	93.5	15.7	713	1	GACL_HUMAN	O75325 homo sapien
14	93	15.6	1018	1	CONF_HUMAN	Q12860 homo sapien
15	92	15.5	702	1	CEAS_HUMAN	P06731 homo sapien
16	91	15.3	555	1	CL66_CARAU	Q90304 carassius a
17	91	15.3	739	1	VCAI_RAT	P29534 rattus norv
18	90.5	15.2	1914	1	KMIS_HUMAN	Q15746 homo sapien
19	89	15.0	1020	1	CONF_MOUSE	P12960 mus musculu
20	87	14.6	3707	1	PGBM_MOUSE	Q05793 mus musculu
21	86.5	14.5	1266	1	NGCA_CHICK	Q03696 gallus gall
22	86	14.5	1021	1	CONF_RAT	Q63198 rattus norv
23	85.5	14.4	4393	1	PGBM_HUMAN	P98160 homo sapien
24	85	14.3	1010	1	CONF_CHICK	P14781 gallus gall
25	85	14.3	1036	1	AXOL_CHICK	P28685 gallus gall
26	84.5	14.2	337	1	G55A_CHICK	Q98892 gallus gall
27	84	14.1	873	1	FAS2_DROME	P34082 drosophila
28	84	14.1	1356	1	VGR2_HUMAN	P35968 homo sapien
29	82.5	13.9	338	1	LAMP_RAT	Q62813 rattus norv
30	82.5	13.9	345	1	OPCM_HUMAN	Q14982 homo sapien
31	82.5	13.9	345	1	OPCM_RAT	P12736 rattus norv
32	82	13.8	673	1	SM4B_HUMAN	Q9NPR2 homo sapien
33	82	13.8	850	1	NRG2_HUMAN	Q14511 homo sapien

34	81.5	13.7	345	1	OPCM_BOVIN	P11834 bos taurus
35	81.5	13.7	782	1	SM4B_MOUSE	Q62179 mus musculu
36	81	13.6	898	1	FAS2_SCHAM	P22648 schistocerc
37	80.5	13.5	338	1	LAMP_HUMAN	Q13449 homo sapien
38	80.5	13.5	1336	1	VGR1_RAT	P53767 rattus norv
39	79	13.3	419	1	PSG1_HUMAN	P11464 homo sapien
40	78.5	13.2	338	1	LAMP_CHICK	Q98919 gallus gall
41	78.5	13.2	756	1	NRG2_MOUSE	P56974 mus musculu
42	78.5	13.2	868	1	NRG2_RAT	O35569 rattus norv
43	78.5	13.2	999	1	MERK_HUMAN	O12866 homo sapien
44	78.5	13.2	1333	1	VGR1_MOUSE	P35969 mus musculu
45	78	13.1	994	1	MERK_RAT	P57097 rattus norv

# ALIGNMENTS

RESULT 1  
RAGE\_HUMAN STANDARD: PPT: 404 AA.  
AC Q15109; Q15279; Q9Y3R3; Q9H2X7;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Advanced glycosylation and product-specific receptor precursor  
DE (Receptor for advanced glycosylation end products)  
DE AGER OR RAGE.  
CN Homo sapiens (Human).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Lung;  
RX MEDLINE=92340547; PubMed=1378843;  
RA Neepser M., Schmidt A.M., Brett J., Yan S.D., Wang F., Pan Y.C.,  
RA Elliston K., Stern D., Shaw A.;  
RT Cloning and expression of a cell surface receptor for advanced  
RT glycosylation end products of proteins.\*;  
RL J. Biol. Chem. 267:14998-15004(1992).

[2]  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=95137587; PubMed=7835890;  
RA Suqaya K., Fukagawa T., Matsumoto K., Mita K., Takahashi E., Ando A.,  
RA Inoko H., Ikemura T.;  
RT \*Three genes in the human MHC class III region near the junction with  
RT the class II: gene for receptor of advanced glycosylation end  
RT products, PBX2 homeobox gene and a notch homolog, human counterpart  
RT of mouse mammary tumor gene int-3.\*;  
RL Genomics 23:408-419(1994).

[3]  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RA Rowen L., Dankers C., Paskin D., Faust J., Loretz C., Ahearn M.E.,  
RA Banta A., Spies T., Hood L.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT APC-100.  
RA Abedin M.J., Yonekura H., Migita H., Karasawa J., Yamamoto Y.,  
RA Yamamoto H.;  
RT \*Molecular heterogeneity of the receptor for advanced glycation  
RT endproducts.\*;  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

[5]  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RA Malherbe P., Richards J., Gaillard H., Thompson A., Diener C.,  
RA Schuler A., Huber G.;  
RT \*cDNA cloning of a novel secreted isoform of the human Receptor for  
RT advanced glycation end products (PAGE) and characterization of cells  
RT co-expressing cell-surface scavenger receptors and Swedish mutant  
RT amyloid precursor protein.\*;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A. (ISOFORM 1).

TISSUE: Lung;  
RA Stansberry R.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/TrEMBL databases.  
RN [1]  
RA Hudson B.L., Futers T.S.;  
RT "Novel polymorphisms in the receptor for advanced glycation  
end-products (RAGE) gene";  
RL Submitted (NOV-1999) to the EMBL/GenBank/TrEMBL databases.  
CC \* FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END  
PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS  
WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED  
RATE IN DIABETES.  
CC \* SUBCELLULAR LOCATION: Type I membrane protein (Isoform 1).  
CC \* Secreted (Isoform 2).  
CC \* ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2/RAGESEC;  
CC ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC \* TISSUE SPECIFICITY: ENDOTHELIAL CELLS.  
CC \* SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
CC \* SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.  
CC \* SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
CC This SWISS PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: M91211; AAA04574.1;  
DR EMBL: D28769; BAA05958.1;  
DR EMBL: U89436; AAB47491.1;  
DR EMBL: AB036442; BAA89469.1;  
DR EMBL: AJ133822; CAB43198.1;  
DR EMBL: BC020669; AAL20669.1;  
DR EMBL: AF208289; AAG35748.1;  
DR Genbank: BCNC3320; AGER.  
DR MIM: 600214;  
DR InterPro: IPR004006; Ig\_MHC.  
DR InterPro: IPR003598; Ig\_C2.  
DR InterPro: IPR003600; Ig\_Like.  
DR Pfam: PF00047; Ig\_2.  
DR SMART: SM00410; Ig\_Like; 1.  
DR SMART: SM00408; IgC2; 1.  
DR PROSITE: PS00290; Ig\_MHC; 1.  
DR Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal;  
KW Alternative splicing; Polymorphism.  
FT SIGNAL 1 22  
FT CHAIN 23 404  
FT POTENTIAL  
FT ADVANCED GLYCOSYLATION END PRODUCT-  
FT SPECIFIC RECEPTOR.  
FT EXTRACELLULAR (POTENTIAL).  
FT POTENTIAL.  
FT CYTOPLASMIC (POTENTIAL).  
FT IG-LIKE V-TYPE DOMAIN.  
FT IG-LIKE C2-TYPE DOMAIN 1.  
FT IG-LIKE C2-TYPE DOMAIN 2.  
FT POTENTIAL.  
FT POTENTIAL.  
FT POTENTIAL.  
FT POTENTIAL.  
FT N LINKED (GLCNAC...) (POTENTIAL).  
FT N LINKED (GLCNAC...) (POTENTIAL).  
FT POLY GLU.  
FT MISSING (IN ISOFORM 2).  
FT GVPLPPLPVLLPPIGPDGCTVSCVAFHSHGPOESKA  
FT VLSLLEPQEEPRTAGVSGSHGLALGLGLGLSTAAL  
FT IGVVILWQPPQPPFEWKAPFNQFEFFFAELNQSPEEG  
FT ESTSGEP -> VSDLEPAGPPTPGAN'RLNRPAPGNS  
FT PQCDQQRCDSPAHWCHIVAKAAVPRKCRDPRKPGQG  
FT CARKLESVGGT (IN ISOFORM 2).  
FT V...R.  
FT /FTID VAR\_011330.  
FT M...G (IN REF. 1).  
FT CONFLICT 1 1

SEQ SEQUENCE 404 AA: 42802 MW. 6058404 460 00VE7 CR644;  
Query Match 100.0%; Score 595; DB 1; Length 404;  
Best local Similarity 100.0%; Pred. NO. 6,9e 55;  
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGNITAPIGEPLVLRCKGAPKPPKLEWKLNTGTEAKKVLSPGGSGHWVAHVLENG 60  
LE 23 AUNTITAPIGEPLVLRCKGAPKPPKLEWKLNTGTEAKKVLSPGGSGHWVAHVLENG 82  
QY 61 SLFLPAVGLQDGRIFPCVAMNNGKETSNSYRVRYVQIGKLEIVISASELT 112  
DB 83 SLFLPAVGLQDGRIFPCVAMNNGKETSNSYRVRYVQIGKLEIVISASELT 144  
RESULT 2  
RAGE\_HOVIN  
ID RAGE\_HOVIN STANDARD; PRT: 416 AA.  
AC Q28173;  
DT 01-NOV-1997 (Ref. 45, Created)  
DI 01-NOV-1997 (Ref. 45, Last sequence update)  
DI 15-JUN-2002 (Ref. 41, Last annotation update)  
DE Advanced glycosylation end product specific receptor precursor  
GN (receptor for advanced glycosylation end products)  
OS AGER OR RAGE.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovina; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OC NCBI\_TaxID:9913;  
OX [1]  
RN [1]  
RE SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
RC TISSUE=Lung;  
RX MEDLINE=92440547; PubMed=1478843;  
RA Neepet M., Schmidt A.M., Brett J., Yan S.D., Wang Y.C.,  
RA Elliston K., Stern D., Shaw A.;  
RA Cloning and expression of a cell surface receptor for advanced  
glycosylation end products of glycosylated proteins.\*  
RL J. Biol. Chem. 267:14998-15004(1992).  
CC \* FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END  
PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS  
WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED  
RATE IN DIABETES.  
CC \* SUBCELLULAR LOCATION: Type I membrane protein  
CC \* TISSUE SPECIFICITY: ENDOTHELIAL CELLS.  
CC \* SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
CC \* SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN  
CC \* SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: M91212; AAA04575.1;  
DR InterPro: IPR004006; Ig\_MHC.  
DR InterPro: IPR003598; Ig\_C2.  
DR InterPro: IPR003600; Ig\_Like.  
DR Pfam: PF00047; Ig\_2.  
DR SMART: SM00410; Ig\_Like; 1.  
DR SMART: SM00408; IgC2; 1.  
DR PROSITE: PS00290; Ig\_MHC; 1.  
DR Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal;  
FT SIGNAL 1 22  
FT CHAIN 23 416  
FT ADVANCED GLYCOSYLATION END PRODUCT-  
FT SPECIFIC RECEPTOR.  
FT EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM  
FT POTENTIAL.  
FT CYTOPLASMIC (POTENTIAL).  
FT IG-LIKE V-TYPE DOMAIN.  
FT IG-LIKE C2-TYPE DOMAIN 1.  
FT IG-LIKE C2-TYPE DOMAIN 2.  
FT POTENTIAL.  
FT POTENTIAL.  
FT POTENTIAL.  
FT POTENTIAL.  
FT N LINKED (GLCNAC...) (POTENTIAL).  
FT N LINKED (GLCNAC...) (POTENTIAL).  
FT POLY GLU.  
FT MISSING (IN ISOFORM 2).  
FT GVPLPPLPVLLPPIGPDGCTVSCVAFHSHGPOESKA  
FT VLSLLEPQEEPRTAGVSGSHGLALGLGLGLSTAAL  
FT IGVVILWQPPQPPFEWKAPFNQFEFFFAELNQSPEEG  
FT ESTSGEP -> VSDLEPAGPPTPGAN'RLNRPAPGNS  
FT PQCDQQRCDSPAHWCHIVAKAAVPRKCRDPRKPGQG  
FT CARKLESVGGT (IN ISOFORM 2).  
FT V...R.  
FT /FTID VAR\_011330.  
FT M...G (IN REF. 1).  
FT CONFLICT 1 1

FT DOMAIN 136 214 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN 262 318 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DISULFID 38 98 POTENTIAL.  
 FT DISULFID 143 207 POTENTIAL.  
 FT DISULFID 269 311 POTENTIAL.  
 FT CARBOHYD 25 25 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 80 80 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT DOMAIN 391 396 POLY-GLU.  
 SQ SEQUENCE 416 AA; 44182 MW; 470381557347677AE CRC64;

Query Match 86 1%; Score 513.5; DB 1; Length 416;  
 Best Local Similarity 89 1%; Pred. No. 2.4e-46;  
 Matches 98; Conservative 5; Mismatches 6; Indels 1; Gaps 1;

QY 2 QNITARIGELPLVLCCKGAPKKPPORLEWKLNTGRTHAKVLSPOGGPWDSVARVLPNGS 61  
 DB 24 QNITARIGELPLVLCCKGAPKKPPORLEWKLNTGRTHAKVLSPOGGPWDSVARVLPNGS 82

QY 62 LFLPAVILQDEGTHFFQVAMNPNKTKTSNYFVYVQIPGKPEIVDSASEL 111  
 DB 83 LLLPAVIGLVDGTHFFQVAMNPNKTKTSNYFVYVQIPGKPEIVDSASEL 132

RESULT 3  
 RAGE\_MOUSE  
 ID PAGE\_MOUSE STANDARD: PRT: 403 AA.  
 AC Q62151;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Advanced glycosylation end product-specific receptor precursor  
 DE (Receptor for advanced glycosylation end products).  
 GN AGPR OR RAGE.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SOURCE FROM N.A.  
 RC STRAIN-BALB/C; TISSUE=Lung;  
 RX MEDLINE=97368045; PubMed=9224812;  
 PA Penard C., Chappay O., Wautier J.M., Nagashima M., Lundh E.,  
 PA Morser J., Zhao L., Schmidt A.M., Schermann J.M., Wautier J.L.,  
 RT "Recombinant advanced glycation end product receptor pharmacokinetics  
 RT in normal and diabetic rats".  
 RL Mol. Pharmacol. 52:54-62(1997).  
 CC -!- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END  
 CC PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS  
 CC WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED  
 CC RATE IN DIABETES.  
 CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.  
 CC -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY  
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: L33412; AAA00040.1;  
 DR MGD: MGI:893592; Ager.  
 DR InterPro: IPR003006; Iq\_MHC.  
 DR InterPro: IPR003598; Iq\_C2.  
 DR InterPro: IPR003600; Iq\_Like.  
 DR Pfam: PF00047; Iq\_3.  
 DR SMART: SM00410; Iq\_Like; 1.  
 DR SMART: SM00408; Iq\_2; 1.  
 DR PROSITE: PS00290; Iq\_MHC; 1.

KW Immunoglobulin domain, Glycoprotein, Transmembrane, Repeat, Signal.  
 FT SIGNAL 1 32 POTENTIAL.  
 FT CHAIN 23 403 ADVANCED GLYCOSYLATION END PRODUCT-SPECIFIC RECEPTOR.  
 FT DOMAIN 23 341 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 342 362 POTENTIAL.  
 FT DOMAIN 363 403 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 31 105 IG-LIKE V-TYPE DOMAIN.  
 FT DOMAIN 136 214 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN 251 307 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DISULFID 38 98 POTENTIAL.  
 FT DISULFID 143 207 POTENTIAL.  
 FT DISULFID 258 300 POTENTIAL.  
 FT CARBOHYD 25 25 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD 80 80 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 403 AA; 42668 MW; 1279796FD1579357 CRC64;

Query Match 86 1%; Score 512.5; DB 1; Length 403;  
 Best Local Similarity 87 4%; Pred. No. 2.9e-46;  
 Matches 97; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 2 QNITARIGELPLVLCCKGAPKKPPORLEWKLNTGRTHAKVLSPOGGPWDSVARVLPNGS 61  
 DB 24 QNITARIGELPLVLCCKGAPKKPPORLEWKLNTGRTHAKVLSPOGGPWDSVARVLPNGS 82

QY 62 LFLPAVILQDEGTHFFQVAMNPNKTKTSNYFVYVQIPGKPEIVDSASEL 112  
 DB 83 LLLPAVIGLVDGTHFFQVAMNPNKTKTSNYFVYVQIPGKPEIVDSASEL 133

RESULT 4  
 RAGE\_RAT  
 ID PAGE\_RAT STANDARD: PRT: 402 AA.  
 AC Q63495;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Advanced glycosylation end product-specific receptor precursor  
 DE (Receptor for advanced glycosylation end products).  
 GN AGER OR RAGE.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SOURCE FROM N.A.  
 RC STRAIN-Sprague-Dawley; TISSUE=Lung;  
 RX MEDLINE=97368045; PubMed=9224812;  
 PA Penard C., Chappay O., Wautier J.M., Nagashima M., Lundh E.,  
 PA Morser J., Zhao L., Schmidt A.M., Schermann J.M., Wautier J.L.,  
 RT "Recombinant advanced glycation end product receptor pharmacokinetics  
 RT in normal and diabetic rats".  
 RL Mol. Pharmacol. 52:54-62(1997).  
 CC -!- FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END  
 CC PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS  
 CC WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED  
 CC RATE IN DIABETES.  
 CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.  
 CC -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN  
 CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: L33413; AAA42027.1;  
 DR InterPro: IPR003006; Iq\_MHC.

```

DR InterPro: IPR004598; Iq_c2.
DR InterPro: IPR003600; Iq_1like.
DR Pfam: PF00047; Iq; 3.
DR SMART: SM00410; Iq_1like; 1.
DR SMART: SM00408; Iq_2; 1.
DR PROSITE: PS00290; Iq_MHC; 1.
KW Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 402
FT DOMAIN 23 441
FT TRANSMEM 342 402
FT DOMAIN 363 402
FT DOMAIN 31 105
FT DOMAIN 136 213
FT DOMAIN 250 306
FT DOMAIN 48 98
FT DISULFID 143 206
FT DISULFID 257 299
FT CARBOHYD 25 25
FT CARBOHYD 80 80
SQ SEQUENCE 402 AA; 42663 MW; 594481BCA51E94FC CRC64;

Query Match 81.1%; Score 494.5; DB 1; Length 402;
Best Local Similarity 84.7%; Pred No. 2 2n-44;
Matches 94; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 2 QNITARIGEPVLEKFKAPKPPDPLEWKTGTEAKVLSPOGPGWDSVARVLNCS 61
DB 24 QNITARIGEPVLEKFKAPKPPDPLEWKTGTEAKVLSPOGPGWDSVARVLNCS 82
QY 62 LFLPAGVIGDEGFEFGDAMNHNKPKSKNYRVVQVIGKDEIVDSASELT 112
DB 83 LFLPAGVIGDEGFEFGDAMNHNKPKSKNYRVVQVIGKDEIVDSASELT 133
QY 83 LFLPAGVIGDEGFEFGDAMNHNKPKSKNYRVVQVIGKDEIVDSASELT 133
DB 83 LFLPAGVIGDEGFEFGDAMNHNKPKSKNYRVVQVIGKDEIVDSASELT 133

RESULT 5
DCC_MOUSE
ID DCC_MOUSE STANDARD PRT: 1447 AA
AC P70211;
DI 01-NOV-1997 (Rel. 35, created)
DI 15 JUN 2002 (Rel. 41, last sequence update)
DI 15 JUN 2002 (Rel. 41, last annotation update)
DE tumor suppressor protein DCC precursor.
GN DCC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RC STRAIN-BALB/C; TISSUE-Brain;
RX MEDLINE=96112625; PubMed=94570174;
RA Cooper H.M., Arnes P., Britto J., Gad J., Wilks A.P.;
RT "Cloning of the mouse homologue of the deleted in colorectal cancer
RL gene (mDCC) and its expression in the developing mouse embryo.";
RL Oncogene 11:2243-2254 (1996).
RN [2]
RP REVISIONS.
RC STRAIN-BALB/C; TISSUE-Brain;
RA Cooper H.M.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: IMPLICATED AS A TUMOR SUPPRESSOR GENE.
CC 1- SUBCELLULAR LOCATION: Type I membrane protein.
CC 1- ALTERNATIVE PRODUCTS: 2 isoforms; A (shown here) and B; are
CC produced by alternative splicing.
CC 1- TISSUE SPECIFICITY: IN THE EMBRYO, EXPRESSED AT HIGH LEVELS IN THE
CC DEVELOPING BRAIN AND NEURAL TUBE. IN ADULT, HIGHLY EXPRESSED IN
CC BRAIN WITH VERY LOW LEVELS FOUND IN TESTIS, HEART AND THYMUS.
CC ISOFORM C IS EXPRESSED ONLY IN THE EMBRYO.
CC 1- DEVELOPMENTAL STAGE: LOW LEVELS IN EARLY GESTATION. HIGHEST LEVELS
CC EXPRESSED DURING MID GESTATION. LEVELS DECREASE IN LATE GESTATION

```

```

AND REMAIN AT THIS LEVEL IN THE ADULT.
-1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. DCC
CC SUBFAMILY.
-1- SIMILARITY: CONTAINS 4 IMMUNOGLOBULIN LIKE C2 TYPE DOMAINS
CC
-1- SIMILARITY: CONTAINS 6 FIBRONECTIN TYPE III LIKE DOMAINS.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/submit/submit.html)
CC or send an email to license@ebi.ac.uk.
CC
CC EMBL: X85788; CAA59786.1;
DR HSSP: P56276; ITLK.
DR MGD: MGI:94869; DCC.
DR InterPro: IPR004961; FN_III.
DR InterPro: IPR004962; FNIII_repeat.
DR InterPro: IPR004006; Iq_MHC.
DR InterPro: IPR004598; Iq_1.
DR InterPro: IPR004600; Iq_2.
DR Pfam: PF00041; Iq; 3.
DR Pfam: PF00047; Iq; 4.
DR PRINTS: PR00014; FNIIIpeat.
DR SMART: SM00410; Iq_1like; 2.
DR SMART: SM00408; Iq_2; 1.
DR SMART: SM00408; Iq_2; 1.
KW Glycoprotein; Immunoglobulin; domain; Transmembrane; Signal.
KW Repeat; Anti-oncogene; Alternative initiation; Alternative splicing.
FT SIGNAL 1 25
FT CHAIN 26 1447
FT INIT MET 85 1447
FT DOMAIN 26 1097
FT TRANSMEM 1098 1122
FT DOMAIN 1123 1447
FT DOMAIN 54 124
FT DOMAIN 154 219
FT DOMAIN 254 317
FT DOMAIN 345 407
FT DOMAIN 426 522
FT DOMAIN 525 618
FT DOMAIN 619 716
FT DOMAIN 722 816
FT DOMAIN 840 940
FT DOMAIN 941 1042
FT DISULFID 61 117
FT DISULFID 161 212
FT DISULFID 261 310
FT CARBOHYD 60 60
FT CARBOHYD 94 94
FT CARBOHYD 299 299
FT CARBOHYD 318 318
FT CARBOHYD 478 478
FT CARBOHYD 628 628
FT CARBOHYD 732 732
FT CARBOHYD 819 838
FT VARSPLIC 1447 AA; 158298 MW; 0D1F1097C2D5B9FC CRC64;
SQ SEQUENCE 1447 AA; 158298 MW; 0D1F1097C2D5B9FC CRC64;

Query Match 19.3%; Score 115; DB 1; Length 1447;
Best Local Similarity 31.3%; Pred. No. 0.00046;
Matches 31; Conservative 20; Mismatches 48; Indels 10; Gaps 3;

QY 2 QNITARIGEPVLEKFKAPKPPDPLEWKTGTEAKVLSPOGPGWDSVARVLNCS 61
DB 147 ESTAFACDVLKKE-VIGEPMTTHWQNG 196
QY 62 LFLPAGVIGDEGFEFGDAMNHNKPKSKNYRVVQVIGKDEIVDSASELT 100
DB 197 LQISRIQPGDSGVKRSARNIVASIRTONIAEVLISDVG 235

```



```

DB 147 EKVIAFMGDIIVLKKE VIGREMPITRWKKNQ - QDLPI - PDSRVVVILPSGA 196
QY 62 LFLPAVGIGDEIEWQAMNPNKFTKSNYVAVYQIPK 100
DB 147 LQISRLQPGDITGVCSARNPASSRTGNEAEVRILSDPG 245

RESULT 7
KMLS_RABIT
ID KMLS_RABIT STANDARD: PRT: 1147 AA.
DI 029294: Q28729;
DI 01-DEC-1992 (Rel. 24, Created)
DI 15 JUN-1998 (Rel. 36, Last sequence update)
DI 16 OCT-2001 (Rel. 40, Last annotation update)
DE MYOSIN light chain kinase, smooth muscle (EC 2.7.1.117) (MLCK)
DE [Contains: Telokin]
GN MYLK
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID-9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE: Smooth muscle;
RX MEDLINE-92084694; PubMed-17486667;
RA Gallagher P.J., Herring B.P., Griffin S.A., Stull T.T.;
RT "Molecular characterization of a mammalian smooth muscle myosin light
chain kinase.";
RL J. Biol. Chem. 266:21946-21944 (1991);
RN [2]
RP SEQUENCE OF 993-1147 FROM N.A. (TELOKIN).
RX MEDLINE-92084695; PubMed-1748667;
RA Gallagher P.J., Herring B.P.;
RT "The carboxyl terminus of the smooth muscle myosin light chain kinase
is expressed as an independent protein, telokin.";
RL J. Biol. Chem. 266:21945-21953 (1991);
CC 1- FUNCTION: PHOSPHORYLATES A SPECIFIC SERINE IN THE N-TERMINUS OF A
MYOSIN LIGHT CHAIN.
CC 1- CATALYTIC ACTIVITY: ATP + [myosin light chain] -> ADP + [myosin
light chain] phosphate.
CC 1- ALTERNATIVE PRODUCTS: AT LEAST THREE ISOFORMS OF MLCK ARE PRODUCED
BY ALTERNATIVE INITIATION, A NON MUSCLE FORM (WHICH IS THE LONGEST
FORM); A SMOOTH-MUSCLE FORM AND TELOKIN (A C-TERMINAL SECTION WITH
NO CATALYTIC ACTIVITY).
CC 1- DOMAIN: TELOKIN BINDS CALMODULIN.
CC 1- SIMILARITY: BELONGS TO THE SPR/THP FAMILY OF PROTEIN KINASES.
CC 1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC 1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
with the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed, usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcements/
or send an email to license@isb-sib.ch).
```

```

DR SMART: SM00060; FN3: 1;
DR SMART: SM00410; IG 1146; 1;
DR SMART: SM00408; IG2: 2;
DR SMART: SM00220; S TK: 1;
DR PROSITE: PS00107; PROTEIN_KINASE_ATP: 1;
DR PROSITE: PS00108; PROTEIN_KINASE_SU: 1;
DR PROSITE: PS00111; PROTEIN_KINASE_DOM: 1;
KW Transferase; Serine/threonine-protein kinase; Calmodulin binding;
KW ATP binding; Phosphorylation; Immunoglobulin domain; Repeat;
FT CHAIN 1 1147
FT CHAIN 1 1147
FT CHAIN 993 1147
FT INL_MET 993 993
FT DOMAIN 103 288
FT REPEAT 100 111
FT REPEAT 112 123
FT REPEAT 124 144
FT REPEAT 133 144
FT REPEAT 145 156
FT REPEAT 157 168
FT REPEAT 169 180
FT REPEAT 181 192
FT REPEAT 193 204
FT REPEAT 205 216
FT REPEAT 217 228
FT REPEAT 229 240
FT REPEAT 243 252
FT REPEAT 253 264
FT REPEAT 265 276
FT REPEAT 277 288
FT DOMAIN 574 644
FT DOMAIN 696 951
FT DOMAIN 943 1006
FT DOMAIN 1056 1124
FT N_BIND 702 710
FT BINDING 725 725
FT ACT_SITE 817 817
FT DOMAIN 1138 1147
FT C_NELCI 1114 1114
; SQ SEQUENCE 1147 AA; F039E62406E1024 -R014;
Query Match 17.3%; Score 103; DR 1; Length 1147;
Best Local Similarity 27.3%; Pred. No. 0.0064;
Matches 40; Conservative 21; Mismatches 39; Indels 20; Gaps 5;
QY 2 QNITARIGEPVLKCKGAPKPPKLEWKLN -TGKTEAWKVLSPGKGGWQSVARVLPN 59
DB 336 QDVRVAREGKHLIQCO-VSSEPPATITWLNCKTKTKTKFVLISQVGSICVSTFKALP- 493
QY 60 GSLFLPAVGIQDEIEWQAMNPNKFTKSNYVAVYQIPK -PEI 104
DB 394 -----EDKGLYK;VAKNA APEAEKSHIVTVIRKAPA;ENAKAFEM 432
RESULT 8
KMLS_BOVIN
ID KMLS_BOVIN STANDARD: PRT: 1176 AA.
DI Q28824;
DI 15 JUN-1998 (Rel. 36, Created)
DI 15 JUN-1998 (Rel. 36, Last sequence update)
DI 16 OCT-2001 (Rel. 40, Last annotation update)
DE Myosin light chain kinase, smooth muscle (EC 2.7.1.117) (MLCK)
DE [Contains: Telokin]
GN MYLK
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovina; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID-9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE: Stomach;
```







[9] SEQUENCE OF 1030-1257 FROM N.A.  
 RX MEDLINE-91132183; PubMed-1993895;  
 RA Harper J.P., Prince J.T., Healy P.A., Nauman S.J.,  
 RA Stallcup W.B.;  
 RT "Isolation and sequence of partial cDNA clones of human L1: homology  
 RT of human and rodent L1 in the cytoplasmic region.";  
 RL J. Neurochem. 56:797-804(1991).  
 [10]  
 RP SEQUENCE OF 20-26.  
 RX MEDLINE-88298876; PubMed-3136158;  
 RA Wolff J.M., Frank R., Mujoo K., Spiro K.C., Reitsfeld R.A.,  
 RA Rathjen F.G.;  
 RT "A human brain glycoprotein related to the mouse cell adhesion  
 RT molecule L1.";  
 RL J. Biol. Chem. 263:11943-11947(1988).  
 [11]  
 RP VAPILANT HSAS TNR-264.  
 RX MEDLINE-94004956; PubMed-8401576;  
 RA Jouet M., Rosenthal A., Macfarlane J., Kenwright S., Donnai D.;  
 RA "A missense mutation confirms the L1 defect in X-linked hydrocephalus  
 RT (HSAS).";  
 RL Nat. Genet. 4:331-331(1993).  
 [12]  
 RP VARIANT HSAS/MASA LEU-1194.  
 RX MEDLINE-95181712; PubMed-7881431;  
 RA Fransen E., Schrander-Stumpel C., Vits L., Coucke P., van Camp G.,  
 RA Willems P.J.;  
 RT "X-linked hydrocephalus and MASA syndrome present in one family are  
 RT due to a single missense mutation in exon 28 of the L1CAM gene.";  
 RL Hum. Mol. Genet. 3:2255-2256(1994).  
 [13]  
 RP VARIANTS HSAS GLN-184 AND ARG-452, AND VARIANT MASA GLN-210.  
 RX MEDLINE-95004608; PubMed-7920659;  
 RA Jouet M., Rosenthal A., Armstrong G., Macfarlane J., Stevenson R.,  
 RA Paterson J., Metzberg A., Ionescu V., Temple K., Kenwright S.;  
 RT "X-linked spastic paraplegia (SPGL), MASA syndrome and X-linked  
 RT hydrocephalus result from mutations in the L1 gene.";  
 RL Nat. Genet. 7:402-407(1994).  
 [14]  
 RP VARIANTS MASA GLN-210 AND ASN 598.  
 RX MEDLINE-95004609; PubMed-7920660;  
 RA Vits L., van Camp G., Coucke P., Fransen E., de Boule K.,  
 RA Heyniers E., Korn B., Boustka A., Wilson G., Schrander-Stumpel C.,  
 RA Winter R.M., Schwartz C., Willems P.J.;  
 RT "MASA syndrome is due to mutations in the neural cell adhesion gene  
 RT L1CAM.";  
 RL Nat. Genet. 7:408-413(1994).  
 [15]  
 RP VARIANTS HSAS/MASA S-9; S-121; K-309, F-768; U-94; AND C-1070.  
 RX MEDLINE-95282776; PubMed-7762552;  
 RA Jouet M., Moncla A., Paterson J., McKewen C., Fryer A., Carpenter N.,  
 RA Holmberg E., Wadelius G., Kenwright S.;  
 RT "New domains of neural cell adhesion molecule L1 implicated in  
 RT X-linked hydrocephalus and MASA syndrome.";  
 RL Am. J. Hum. Genet. 56:1304-1314(1995).  
 [16]  
 RP VARIANTS HSAS/MASA Q-184; Q-210; Y-264; R-452; N-598 AND L-1194.  
 RX MEDLINE-96153146; PubMed-8556302;  
 RA Fransen E., Lemmon V., van Camp G., Vits L., Coucke P., Willems P.J.;  
 RA "CRASH syndrome: clinical spectrum of corpus callosum hypoplasia,  
 RT retardation, adducted thumbs, spastic paraparesis and hydrocephalus  
 RT due to mutations in one single gene, L1.";  
 RL Eur. J. Hum. Genet. 3:273-284(1995).  
 [17]  
 RP KRRATUM.  
 RA Fransen E., Lemmon V., van Camp G., Vits L., Coucke P., Willems P.J.;  
 RL Eur. J. Hum. Genet. 4:126-126(1996).  
 [18]  
 RP VARIANTS HSAS/MASA/SPG1 SER-179 AND ARG-370.  
 RX MEDLINE-96057511; PubMed-7562969;  
 RA Ruiz J.C., Cuppens H., Legius E., Fryns J.-P., Glover T., Marynen P.,  
 RA Cassiman J.-J.;

"Mutations in L1-CAM in two families with X linked complicated  
 RT spastic paraplegia, MASA syndrome, and HSAS";  
 RL J. Med. Genet. 32:549-552(1995).  
 [19]  
 RP VARIANTS HSAS CYS-194 AND LEU-240.  
 RX MEDLINE-97083370; PubMed-8929944;  
 RA Gu S.-M., Orth U., Veske A., Enders H., Kluender K., Schloesser M.,  
 RA Engel W., Schwinger E., Gal A.;  
 RT "Five novel mutations in the L1CAM gene in families with X linked  
 RT hydrocephalus.";  
 RL J. Med. Genet. 33:103-106(1996).  
 [20]  
 RP VARIANTS HSAS Q-184; 439-V-T-443 DEL; C-784 AND 936-G-L-948 DEL.  
 RX MEDLINE-97338664; PubMed-9195224;  
 RA Macfarlane J.R., Du J.-S., Pepys M.E., Ramsden S., Donnai D.,  
 RA Charlton R., Garrett C., Tolmie J., Yates J.R.W., Berry C., Goudie D.,  
 RA Moncla A., Iunt P., Hodgson S., Jouet M., Kenwright S.;  
 RT "Nine novel L1 CAM mutations in families with X-linked  
 RT hydrocephalus.";  
 RL Hum. Mutat. 9:512-518(1997).  
 [21]  
 RP VARIANTS HSAS/MASA ASP-691; ARG-698 AND PRO-935.  
 RX MEDLINE-98180721; PubMed-9521434;  
 RA Du Y.-Z., Srivastava A.K., Schwartz C.E.;  
 RT "Multiple exon screening using restriction endonuclease  
 RT fingerprinting (REF): detection of six novel mutations in the L1 cell  
 RT adhesion molecule (L1CAM) gene.";  
 RL Hum. Mutat. 11:232-230(1998).  
 [22]  
 RP VARIANT CRASH PRO-632.  
 RX MEDLINE-98112489; PubMed-9452110;  
 RA Vits L., Chitayat D., van Camp G., Holden J.J.A., Fransen E.,  
 RA Willems P.J.;  
 RT "Evidence for somatic and germline mosaicism in CRASH syndrome.";  
 RL Hum. Mutat. Suppl. 1:S284-S287(1998).  
 [23]  
 RP VARIANTS HSAS/MASA THR-219; ARG-335; CYS-386; CYS-473 AND LEU-1224.  
 RX MEDLINE-98415726; PubMed-9744477;  
 RA Saugier-Verber P., Martin C., le Meur N., Lyonnet S., Munnich A.,  
 RA David A., Henocq A., Heron D., Jonveaux P., Odent S., Manouvrier S.,  
 RA Moncla A., Moirichon N., Philip N., Salpe D., Tosi M., Frebourg T.;  
 RT "Identification of novel L1CAM mutations using fluorescence-assisted  
 RT mismatch analysis.";  
 RL Hum. Mutat. 13:259-366(1998).  
 [24]  
 RP FUNCTION: CELL ADHESION MOLECULE WITH AN IMPORTANT ROLE IN THE  
 CC DEVELOPMENT OF THE NERVOUS SYSTEM. INVOLVED IN NEURON-NEURON  
 CC ADHESION. NEURITE FASCICULATION, OUTGROWTH OF NEURITES, ETC. BINDS  
 CC TO AXONIN ON NEURONS.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms: 1 (shown here) and 2; are  
 CC produced by alternative splicing.  
 CC -!- DISEASE: DEFECTS IN L1CAM ARE THE CAUSE OF THREE X-LINKED  
 CC SYNDROMES: 1: HYDROCEPHALUS OWING TO STENOSIS OF THE AQUEDUCT OF  
 CC SYLVIVUS (HSAS) CHARACTERIZED BY MENTAL RETARDATION AND ENLARGED

## Query Match

Best Local Similarity 16.2%; Score 96.5; DB 1; Length 1257;

Matches 35; Conservative 12; Mismatches 40; Indels 21; Gaps 4;

QY 1 AONTARIGEPLVLK--KGAKPPQRLKLTGRTAEKVKISPOGGPMDSVARVLP 58

Db 249 SSHLVALCGQPLVLECIAEGFP-----TPTKWLRLPSGMPADRVITYQNH 293

QY 59 NGSFLPAVGTODEGIFRCQAMNNGKETSRYRVYQIP---GRPE 103

Db 294 NKTQLQLKVGEDDGEYRCIAENSLG-SARHAYVTVEAAPYWLHKPQ 340

## RESULT 11

AXOL\_RAT

ID AXOL\_RAT

AC P22063;

DT 01-AUG-1991 (rel. 19, Created)

STANDARD: PRT: 1040 AA.

01 AUG 1991 (Ref.: 1). Last sequence update:  
15-JUN 2002 (Ref.: 4). Last annotation update:  
Contact in z precursor (Axonal-1) (Axonal glycoprotein TAG-1)  
DE (Transient axonal glycoprotein 1) (TAG-1).  
GN CNTN2 OR TAXL.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
NCBI\_TaxID=10116;  
[1]  
RN  
RS SEQUENCE FROM N.A., AND SEQUENCE OF 41 41.  
RT TISSUE=Spinal cord;  
RX MEDLINE=90199890; PubMed=247872;  
RA Furley A.J., Morton S.B., Manalo D., Karanacos D., Dodd J.,  
Jessel T.M.;  
RT "The axonal glycoprotein TAG-1 is an immunoglobulin superfamily  
member with neurite outgrowth-promoting activity";  
Cell 61:157-170(1990).  
RL  
CC 1 FUNCTION: MAY PLAY A ROLE IN THE INITIAL GROWTH AND GUIDANCE OF  
AXONS. MAY BE INVOLVED IN CELL ADHESION.  
CC 1 SUBCELLULAR LOCATION: ATTACHED TO THE NEURONAL MEMBRANE BY A  
GPI ANCHOR AND IS ALSO RELEASED FROM NEURONS.  
CC 1 TISSUE SPECIFICITY: IN NEURAL TISSUES IN EMBRYOS, AND IN ADULT  
BRAIN, SPINAL CORD AND CEREBELLUM.  
CC 1 DEVELOPMENTAL STAGE: TRANSIENTLY EXPRESSED ON A SUBSET OF AXONS  
IN THE DEVELOPING RAT NERVOUS SYSTEM.  
CC 1 SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
CC 1 SIMILARITY: CONTAINS 6 IMMUNOGLOBULIN LIKE C2-TYPE DOMAINS.  
CC 1 SIMILARITY: CONTAINS 4 FIBRONECTIN TYPE III-LIKE DOMAINS.  
CC  
CC This SWISS PROT entry is copyright © it is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb.scripps.edu/>  
or send an email to [license@isb.scripps.edu](mailto:license@isb.scripps.edu))

[illegible]

```

F1 CARGOBYD 920 N LINKED (CLONAL) (PUBLISHED).
F1 CARGOBYD 942 942 N LINKED (CLONAL) (PUBLISHED).
SQ SEQUENCE 1040 AA; 113642 MW; 6E707EE6614CB4B7 CRC64;

Query Match
Best local Similarity 40.5%; Pred. No. 0.941;
Matches 25; Conservative 16; Mismatches 25; Gaps 4;

UY 6 ARIGEPVLVKCKAPKKPPQPLEWKLNTGRTEAKWKVLSFGGGGFWISVARVLPNGSLTLP 65
UR 253 ALVGQGVLECFATSNVPPEPKKPKVGS- LSPQ WATA EFTQLQP 296
UY 66 AVCIQIEGTFRQAMNRNCKET 87
UR 3 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DY 297 SVSFEETGYCEAENSGRDT 418
UR 3 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
UN52.CABEL STANDARD: PRO: 24H1 AA
ID UN52.CABEL Q06561;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Basement membrane proteoglycan precursor (perlecan homodimer)
OS DNC-52;
GN Caenorhabditis elegans.
OR Eukaryota, Metazoa, Nematoda; Chromadorea; Rhabditiada; Rhabditidae;
OC Rhabditiida; Peloderiinae; Caenorhabditis.
RX NCBI TAXID=6239;
LN 111
SEQUENCE FROM N.A.
MEDLINE=94359574; PubMed=849414;
RG Rogalski T.M., Williams B.D., Mullin G.P.; Moerman D.G.
AB "Products of the dnc-52 gene in Caenorhabditis elegans are homologous
to the core protein of the mammalian basement membrane heparan
sulfate proteoglycan."
RI Genes dev. 7:1417-1484(1994).
CC 1 FUNCTION: POSSIBLE ROLE IN MYOFILAMENT ASSEMBLY AND/OR ATTACHMENT
OF THE MYOFILAMENT LATTICE TO THE CELL MEMBRANE. DNC-52 MAY BE AN
EXTRACELLULAR ANCHOR FOR INTEGRIN RECEPTORS IN MUSCLE.
CC 1 SUBCELLULAR LOCATION: Extracellular matrix.
CC 1 ALTERNATIVE PRODUCTS: At least 3 isoforms may be produced by
alternative splicing.
CC 1 TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANE OF ALL
CONTRACTILE TISSUES. IT IS CONCENTRATED OVER MUSCLE DENSE BODIES
AND M-LINES WHICH ARE ASSOCIATED WITH Z-DISK INTEGRIN.
CC 1 DEVELOPMENTAL STAGE: SYNTHESIZED EARLY IN EMBRYOGENESIS.
CC 1 SIMILARITY: CONTAINS 3 LDL RECEPTOR CLASS A DOMAINS.
CC 1 SIMILARITY: CONTAINS 16 IMMUNOGLOBULIN LIKE C2 TYPE DOMAINS.
CC 1 SIMILARITY: CONTAINS 7 LAMININ EGF LIKE DOMAINS.
CC 1 SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.
CC 1 SIMILARITY: TO PERLECAN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities to print a database name at (SwissProt) www.ebi.ac.uk/databases/
or send an email to direct@ebi.ac.uk.

EMBL: L13458; AAA28156.1;
HSSP: P01140; ILLDR
InterPro: IPRO00561; EGF-like.
InterPro: IPRO0006; IQ_MHC.
InterPro: IPRO004598; IQ_C2.
InterPro: IPRO004600; IQ_Like.
InterPro: IPRO02172; LDL_recept_A.
InterPro: IPRO00004; Laminin_B.
InterPro: IPRO02049; Laminin_EGF.
InterPro: IPRO00519; p_trofil.
PIR: PF000047; ba_16;

```

DR Pfam: PF00052; laminin\_B; 2.  
 DR Pfam: PF00053; laminin\_EGF; 5.  
 DR Pfam: PF00057; ldl\_receptor; 3.  
 DR ProDom: PD003031; laminin\_B; 2.  
 DR SMART: SM00180; EGF\_Lam; 5.  
 DR SMART: SM00410; IG\_Like; 5.  
 DR SMART: SM00408; IG2; 11.  
 DR SMART: SM00192; Ldca; 3.  
 DR SMART: SM00281; Lamb; 2.  
 DR SMART: SM00014; P; PARTIAL  
 DR PROSITE: PS00022; EGF\_1; 4.  
 DR PROSITE: PS01186; PCF\_2; 2.  
 DR PROSITE: PS01209; LDLRA\_1; 3.  
 DR PROSITE: PS00068; LDLRA\_2; 3.  
 DR PROSITE: PS01248; LAMININ\_TYE\_EGF; 7.  
 KW Proteoglycan; Glycoprotein; Signal; Immunoglobulin domain; Repeat;  
 KW Basement membrane; Extracellular matrix; Alternative splicing;  
 KW laminin EGF-like domain.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 2481  
 FT DOMAIN 19 148 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN 149 188 LDL-RECEPTOR CLASS A 1.  
 FT DOMAIN 190 229 LDL-RECEPTOR CLASS A 2.  
 FT DOMAIN 233 272 LDL-RECEPTOR CLASS A 3.  
 FT DOMAIN 273 363 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DOMAIN 384 431 LAMININ EGF-LIKE 1 (INCOMPLETE).  
 FT DOMAIN 432 441 LAMININ EGF-LIKE 2 (N-TERMINAL).  
 FT DOMAIN 442 632 LAMININ DOMAIN IV 1.  
 FT DOMAIN 633 665 LAMININ EGF-LIKE 2 (C-TERMINAL).  
 FT DOMAIN 673 719 LAMININ EGF-LIKE 3 (INCOMPLETE).  
 FT DOMAIN 720 729 LAMININ EGF-LIKE 4 (N-TERMINAL).  
 FT DOMAIN 730 920 LAMININ DOMAIN IV 2.  
 FT DOMAIN 921 953 LAMININ EGF-LIKE 4 (C-TERMINAL).  
 FT DOMAIN 954 1003 LAMININ EGF-LIKE 5.  
 FT DOMAIN 1010 1059 LAMININ EGF-LIKE 6.  
 FT DOMAIN 1060 1110 LAMININ EGF-LIKE 7.  
 FT DOMAIN 1133 1216 IG-LIKE C2-TYPE DOMAIN 3.  
 FT DOMAIN 1227 1311 IG-LIKE C2-TYPE DOMAIN 4.  
 FT DOMAIN 1318 1403 IG-LIKE C2-TYPE DOMAIN 5.  
 FT DOMAIN 1415 1499 IG-LIKE C2-TYPE DOMAIN 6.  
 FT DOMAIN 1507 1592 IG-LIKE C2-TYPE DOMAIN 7.  
 FT DOMAIN 1598 1682 IG-LIKE C2-TYPE DOMAIN 8.  
 FT DOMAIN 1699 1785 IG-LIKE C2-TYPE DOMAIN 9.  
 FT DOMAIN 1794 1878 IG-LIKE C2-TYPE DOMAIN 10.  
 FT DOMAIN 1887 1976 IG-LIKE C2-TYPE DOMAIN 11.  
 FT DOMAIN 1977 2070 IG-LIKE C2-TYPE DOMAIN 12.  
 FT DOMAIN 2078 2162 IG-LIKE C2-TYPE DOMAIN 13.  
 FT DOMAIN 2174 2260 IG-LIKE C2-TYPE DOMAIN 14.  
 FT DOMAIN 2267 2347 IG-LIKE C2-TYPE DOMAIN 15.  
 FT DOMAIN 2354 2436 IG-LIKE C2-TYPE DOMAIN 16.  
 FT DISULFID 66 114 BY SIMILARITY.  
 FT DISULFID 149 161 BY SIMILARITY.  
 FT DISULFID 156 174 BY SIMILARITY.  
 FT DISULFID 168 184 BY SIMILARITY.  
 FT DISULFID 190 202 BY SIMILARITY.  
 FT DISULFID 197 215 BY SIMILARITY.  
 FT DISULFID 209 224 BY SIMILARITY.  
 FT DISULFID 233 246 BY SIMILARITY.  
 FT DISULFID 240 259 BY SIMILARITY.  
 FT DISULFID 253 268 BY SIMILARITY.  
 FT DISULFID 954 963 BY SIMILARITY.  
 FT DISULFID 956 970 BY SIMILARITY.  
 FT DISULFID 973 982 BY SIMILARITY.  
 FT DISULFID 985 1001 BY SIMILARITY.  
 FT DISULFID 1010 1020 BY SIMILARITY.  
 FT DISULFID 1012 1026 BY SIMILARITY.  
 FT DISULFID 1029 1038 BY SIMILARITY.  
 FT DISULFID 1041 1057 BY SIMILARITY.  
 FT DISULFID 1060 1068 BY SIMILARITY.  
 FT DISULFID 1062 1078 BY SIMILARITY.  
 FT DISULFID 1081 1090 BY SIMILARITY.  
 FT DISULFID 1093 1108 BY SIMILARITY.  
 FT DISULFID 1151 1199 BY SIMILARITY.  
 PFam: PF00052; laminin\_B; 2.  
 PFam: PF00053; laminin\_EGF; 5.  
 PFam: PF00057; ldl\_receptor; 3.  
 ProDom: PD003031; laminin\_B; 2.  
 SMART: SM00180; EGF\_Lam; 5.  
 SMART: SM00410; IG\_Like; 5.  
 SMART: SM00408; IG2; 11.  
 SMART: SM00192; Ldca; 3.  
 SMART: SM00281; Lamb; 2.  
 SMART: SM00014; P; PARTIAL  
 PROSITE: PS00022; EGF\_1; 4.  
 PROSITE: PS01186; PCF\_2; 2.  
 PROSITE: PS01209; LDLRA\_1; 3.  
 PROSITE: PS00068; LDLRA\_2; 3.  
 PROSITE: PS01248; LAMININ\_TYE\_EGF; 7.  
 Proteoglycan; Glycoprotein; Signal; Immunoglobulin domain; Repeat;  
 Basement membrane; Extracellular matrix; Alternative splicing;  
 laminin EGF-like domain.  
 SIGNAL 1 18 POTENTIAL.  
 CHAIN 19 2481  
 DOMAIN 19 148 IG-LIKE C2-TYPE DOMAIN 1.  
 DOMAIN 149 188 LDL-RECEPTOR CLASS A 1.  
 DOMAIN 190 229 LDL-RECEPTOR CLASS A 2.  
 DOMAIN 233 272 LDL-RECEPTOR CLASS A 3.  
 DOMAIN 273 363 IG-LIKE C2-TYPE DOMAIN 2.  
 DOMAIN 384 431 LAMININ EGF-LIKE 1 (INCOMPLETE).  
 DOMAIN 432 441 LAMININ EGF-LIKE 2 (N-TERMINAL).  
 DOMAIN 442 632 LAMININ DOMAIN IV 1.  
 DOMAIN 633 665 LAMININ EGF-LIKE 2 (C-TERMINAL).  
 DOMAIN 673 719 LAMININ EGF-LIKE 3 (INCOMPLETE).  
 DOMAIN 720 729 LAMININ EGF-LIKE 4 (N-TERMINAL).  
 DOMAIN 730 920 LAMININ DOMAIN IV 2.  
 DOMAIN 921 953 LAMININ EGF-LIKE 4 (C-TERMINAL).  
 DOMAIN 954 1003 LAMININ EGF-LIKE 5.  
 DOMAIN 1010 1059 LAMININ EGF-LIKE 6.  
 DOMAIN 1060 1110 LAMININ EGF-LIKE 7.  
 DOMAIN 1133 1216 IG-LIKE C2-TYPE DOMAIN 3.  
 DOMAIN 1227 1311 IG-LIKE C2-TYPE DOMAIN 4.  
 DOMAIN 1318 1403 IG-LIKE C2-TYPE DOMAIN 5.  
 DOMAIN 1415 1499 IG-LIKE C2-TYPE DOMAIN 6.  
 DOMAIN 1507 1592 IG-LIKE C2-TYPE DOMAIN 7.  
 DOMAIN 1598 1682 IG-LIKE C2-TYPE DOMAIN 8.  
 DOMAIN 1699 1785 IG-LIKE C2-TYPE DOMAIN 9.  
 DOMAIN 1794 1878 IG-LIKE C2-TYPE DOMAIN 10.  
 DOMAIN 1887 1976 IG-LIKE C2-TYPE DOMAIN 11.  
 DOMAIN 1977 2070 IG-LIKE C2-TYPE DOMAIN 12.  
 DOMAIN 2078 2162 IG-LIKE C2-TYPE DOMAIN 13.  
 DOMAIN 2174 2260 IG-LIKE C2-TYPE DOMAIN 14.  
 DOMAIN 2267 2347 IG-LIKE C2-TYPE DOMAIN 15.  
 DOMAIN 2354 2436 IG-LIKE C2-TYPE DOMAIN 16.  
 DISULFID 66 114 BY SIMILARITY.  
 DISULFID 149 161 BY SIMILARITY.  
 DISULFID 156 174 BY SIMILARITY.  
 DISULFID 168 184 BY SIMILARITY.  
 DISULFID 190 202 BY SIMILARITY.  
 DISULFID 197 215 BY SIMILARITY.  
 DISULFID 209 224 BY SIMILARITY.  
 DISULFID 233 246 BY SIMILARITY.  
 DISULFID 240 259 BY SIMILARITY.  
 DISULFID 253 268 BY SIMILARITY.  
 DISULFID 954 963 BY SIMILARITY.  
 DISULFID 956 970 BY SIMILARITY.  
 DISULFID 973 982 BY SIMILARITY.  
 DISULFID 985 1001 BY SIMILARITY.  
 DISULFID 1010 1020 BY SIMILARITY.  
 DISULFID 1012 1026 BY SIMILARITY.  
 DISULFID 1029 1038 BY SIMILARITY.  
 DISULFID 1041 1057 BY SIMILARITY.  
 DISULFID 1060 1068 BY SIMILARITY.  
 DISULFID 1062 1078 BY SIMILARITY.  
 DISULFID 1081 1090 BY SIMILARITY.  
 DISULFID 1093 1108 BY SIMILARITY.  
 DISULFID 1151 1199 BY SIMILARITY.  
 Query Match. 16.1%; Score 95.5; DB 1; Length 2481;  
 Best Local Similarity 25.4%; Pred. No. 0.092;  
 Matches 31; Conservative 21; Mismatches 41; Indels 29; Gaps 5;  
 QY 8 IGEPVLVCKK--GAPKKPPORLEWKLNTGRTAEKWLSPGGG-PWDSVARVLPNGSLFL 64  
 DL 1989 VGEPLVKCEAFGAPGDPEVEWLHDG-----PERGDLDDPKPVTISEQFIRH 2039  
 QY 65 PAVGIODEGIFRCQANRNGKTKSNY-----RVRVYQIDG-----KPEIVDS 107  
 DL 2040 PNVGLNAGVYTCCKSSAHATATKNIYIEVVEPSRIATVSLGSSQWFDOGEKELICT 2099  
 QY 108 AS 109  
 DL 2100 AT 2101  
 RESULT 13  
 GAC1\_HUMAN STANDARD; PRT: 713 AA.  
 AC 075325; 2001 (Rel 40, Created)  
 DT 16-OCT-2001 (Rel 40, Last sequence update)  
 DT 15-JUN-2002 (Rel 41, Last annotation update)  
 DE Glioma amplified on chromosome 1 protein precursor.  
 GN GAC1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Glial tumor;  
 RX MEDLINE-98324709; PubMed-9662332;  
 RA Malfroy B., Almeida A., Zhu X.X., Vogt N., Tyagi R., Muleris M.,  
 RA Dutrillaux A.-M., Dutrillaux B., Ross D., Hanash S.;  
 RT "GAC1, a new member of the leucine-rich repeat superfamily on  
 RT chromosome band 1q32.1, is amplified and overexpressed in malignant  
 RT gliomas.";  
 KL Oncogene 16:2947-3002(1998).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential)  
 CC -!- TISSUE SPECIFICITY: OVERAMPLIFIED IN MALIGNANT GLIOMAS.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 11 LEUCINE-RICH REPEATS (LRR).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 DR EMBL: AF030435; AAC39792.1;  
 DR MIM: 605492;  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003598; Ig\_c2.





```

FT CARBOHYD 309 309 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 330 330 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 351 351 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 360 360 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 375 375 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 432 432 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 466 466 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 480 480 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 508 508 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 529 529 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 553 553 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 560 560 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 580 580 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 612 612 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 650 650 N-LINKED (GLCNAC) (POTENTIAL)
FT CARBOHYD 665 665 N-LINKED (GLCNAC) (POTENTIAL)
FT CONFLICT 420 420 MISSING (IN REF. 4)
SQ SEQUENCE 702 AA: 76795 MW: 6297AE26CDBB85C CRC64:

Query Match 15.5% Score 92; DB 1; Length 702;
Best local Similarity 26.9% Pred. No. 0.054;
Matches 29; Conservative 14; Mismatches 44; Indels 22; Gaps 4;

QY 3 NITATIGFPLIVKCKAPKPKPPPIEWKLNTEIEMKVLSPQGGPPWLSVAVLPNGSL 42
Db 246 NTSYRSGENISCHAA SNPPAQYSWFVN-----GGGTFQQSTQ ---EL 284

QY 63 FLTAVGLQDGIHFCQAMNKNCKETKSNRYRVRYQIIGKPKPIVDSAS 109
Db 285 FIPNITVNNSSGYTQAHNSDTGLNKTTVTITTVYAEPIKPKPITSNNS 332

```

Search completed: May 30, 2004, 15:55:37  
 Job time : 8.55856 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 15:47:46, Search time 25,982 seconds  
(without alignments)  
888,203 million cell updates/sec

Title: US-09-872-185b-1

Perfect score: 595

Sequence: 1 AQLIARIFFPIVIRK-KAP VEVYILPFFPLIVSASLT 112

Scoring table: BLOSUM62

Gapop 10 0 0 Gapext 0 5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21:\*

1: sp-archaea:\*

2: sp-bacteria:\*

3: sp-fungi:\*

4: sp-human:\*

5: sp-invertebrate:\*

6: sp-mammal:\*

7: sp-mbc:\*

8: sp-organelle:\*

9: sp-phage:\*

10: sp-plant:\*

11: sp-rodent:\*

12: sp-virus:\*

13: sp-vertebrate:\*

14: sp-unclassified:\*

15: sp-rvivirus:\*

16: sp-bacterioph:\*

17: sp-archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

# SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	516.5	86.8	402	11	Q35444
2	118	19.8	42	6	Q9TQ01
3	114	19.2	1445	11	Q63155
4	101	17.0	1212	5	Q95T60
5	101	17.0	2016	5	Q9NEA1
6	101	17.0	2016	5	Q9V4J9
7	99	16.6	628	4	Q9TQ01
8	96.5	16.2	1248	6	Q9TQ01
9	96	16.1	1018	5	Q28106
10	95.5	16.1	1344	11	Q92214
11	95.5	16.1	2295	5	Q9XTG3
12	95.5	16.1	2492	5	Q18263
13	95.5	16.1	3375	5	Q9XTG3
14	95	16.0	182	4	Q15402
15	95	16.0	1034	4	Q96MS0
16	95	16.0	1427	13	Q91562

17	94	15.8	1289	4	Q9Y2A5	Q9Y2A5 homo sapien
18	93.5	15.7	756	13	Q8Q609	Q8Q609 gallus gall
19	93.5	15.7	793	11	Q70246	Q70246 mus musculu
20	93.5	15.7	2838	4	Q9NR99	Q9NR99 homo sapien
21	93	15.6	1280	13	Q90933	Q90933 gallus gall
22	93	15.6	1561	11	Q92402	Q92402 mus musculu
23	92.5	15.5	330	13	Q92242	Q92242 gallus gall
24	92.5	15.5	437	4	Q9ULQ7	Q9ULQ7 homo sapien
25	92.5	15.5	1612	11	Q89026	Q89026 mus musculu
26	92.5	15.5	1651	4	Q9Y6N7	Q9Y6N7 homo sapien
27	92.5	15.5	1651	11	Q55005	Q55005 rattus norv
28	91.5	15.4	1389	13	Q90269	Q90269 brachydanio
29	91.5	15.4	1419	13	Q98SW3	Q98SW3 brachydanio
30	91	15.3	626	11	Q92120	Q92120 mus musculu
31	91	15.3	739	11	Q63669	Q63669 rattus norv
32	91	15.3	1040	13	Q9W675	Q9W675 brachydanio
33	91	15.3	1311	5	Q961K8	Q961K8 drosophila
34	91	15.3	1527	5	Q9V224	Q9V224 drosophila
35	91	15.3	1535	5	Q23591	Q23591 drosophila
36	91	15.3	2459	11	Q9JIE1	Q9JIE1 rattus norv
37	91	15.3	5636	4	Q96KW7	Q96KW7 homo sapien
38	90.5	15.2	992	4	Q9C015	Q9C015 homo sapien
39	89	15.0	466	5	Q9S510	Q9S510 drosophila
40	88.5	14.9	333	13	Q90241	Q90241 gallus gall
41	88.5	14.9	1380	4	Q9HCK4	Q9HCK4 homo sapien
42	88.5	14.9	1513	13	Q90270	Q90270 brachydanio
43	88.5	14.9	1614	13	Q8UV07	Q8UV07 xenopus lae
44	88	14.8	213	4	Q8W91	Q8W91 homo sapien
45	88	14.8	332	4	Q15242	Q15242 homo sapien

## ALIGNMENTS

RESULT 1

Q35444 ID Q35444 PPELITINAPY: PPT: 402 AA.  
AC Q35444;

DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)

DE RAGE.

GN RAGE.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

EX NCBI\_taxID=10090;

EN [1]

PP SEQUENCE FROM N.A.

PA Popen L., Mahairas G., Qin S., Ahearn M.E., Dinkers C., Lasky S.,

FA Loretz C., Schmidt S., Tiftos S., Traloff K., Zackroff R., Hood L.;

RT "Sequence of the mouse major histocompatibility locus class III

RT region."

RL Submitted (ref-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF030001; AB082007; Iq\_c2.

DR InterPro: IPR003600; Iq\_c2.

DR InterPro: IPR003600; Iq\_c2.

DR InterPro: IPR003600; Iq\_c2.

DR Pfam: PF00047; Iq; 3.

DR SMART: SM00408; Iq; 1.

DR SMART: SM00410; Iq; 1.

DR ProSITE: PS00240; Iq; 1.

DR Immunoglobulin domain.

SL SEQUENCE: 402 AA; 4255 MW; 108550AUCBCH902 QPC64;

Query Match

Best Local Similarity 88.3%; Score 516.5; DB 11; Length 402;

Mismatches 98; Conserved 3; Mismatches 9; Indels 1; Gaps 1;

Qy 2 QNTARICPLVLCCKAPKKPPQRIEKLINTGRTPEAKVLSPOGGGPDMSVARVLPNGS 61

|||||

24 QNTARICPLVLSCKAPKKPPQRIEKLINTGRTPEAKVLSPOGGGPDMSVARVLPNGS 82





DI 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 15, Last sequence update)  
 DE DSCAM precursor  
 GN DSCAM OR CG17800.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta,  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN 111  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20348742; PubMed=10872553;  
 RA Schmucker D., Clemens J.C., Shu H., Worby C.A., Xiao J., Muda M.,  
 RA Dixon J.E., Zipursky S.L.;  
 RT "Drosophila Dscam is an Axon Guidance Receptor Exhibiting  
 RT Extraordinary Molecular Diversity";  
 RI Cell 101:671-684(2000).  
 DR FBL: AF260530; AAF71926.1; -;  
 DR HSPF: P40189; IBDQ.  
 DR FlyBase; FBgn0033159; Dscam.  
 DR InterPro: IPR000267; Asp/Glutamase.  
 DR InterPro: IPR003961; FN.III.  
 DR InterPro: IPR003598; Ig\_C2.  
 DR InterPro: IPR003600; Ig\_Like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00041; fn3; 6.  
 DR Pfam: PF00047; Ig; 9.  
 DR SMART: SM00060; FN3; 6.  
 DR SMART: SM00408; IgC2; 7.  
 DR SMART: SM00410; Ig\_Like; 3.  
 DR PROSITE: PS00144; ASN\_GLN\_ASE\_1; UNKNOWN\_1.  
 KW Immunoglobulin domain; Signal.  
 FT SIGNAL 1 28 POTENTIAL.  
 SQ SEQUENCE 2016 AA; 222124 MW; 95CE95488F2AD36C CRC64;  
 Query Match 17.0%; Score 101; DB 5; Length 2016;  
 Best Local Similarity 29.2%; Pred. No. 0.06;  
 Matches 33; Conservative 18; Mismatches 48; Indels 14; Gaps 5;  
 QY 2 ONITARIGEPVLVKCGAPKPPORLEWKLNTGRTAEWKVLSPOGGGPPWSVARVLPNG- 60  
 Db 819 RNOTARGEPVAVLQCEAKGEPGIGL-WNNMNM-----LDPKNDNRYTIREEILSTGV 871  
 QY 61 -SLPLPAVGIDEGIPRCQAMNNGKETSRYRVYQIPGKP---EIVDSA 108  
 Db 872 MSSLSIKRTSRSDSALETCTVATNAFGSDAS-INMIVQEVPEMPYALKVLVDKS 923  
 RESULT 6  
 ID Q9V4J9 PRELIMINARY; PRT; 2016 AA.  
 AC Q9V4J9;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE CG17800 protein.  
 GN DSCAM OR CG17800.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN 111  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananidis P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
 RA Brandon P.C., Rogers Y.-H., Blazei R.G., Champe M., Pfeiffer B.B.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,

RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Herman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.R., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evansgill C.C., Ferraz C., Ferrer S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kocira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei P., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Sidon-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong P., Sun E.,  
 RA Svirkas P., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu B., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 PL Science 287:2185-2195(2000).  
 DR FBL: A6003841; AAF59271.1; -;  
 DR HSPF: P40189; IBDQ.  
 DR FlyBase; FBgn0033159; Dscam.  
 DR InterPro: IPR000267; Asp/Glutamase.  
 DR InterPro: IPR003961; FN.III.  
 DR InterPro: IPR003598; Ig\_C2.  
 DR InterPro: IPR003600; Ig\_Like.  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR Pfam: PF00041; fn3; 6.  
 DR Pfam: PF00047; Ig; 9.  
 DR SMART: SM00060; FN3; 6.  
 DR SMART: SM00408; IgC2; 7.  
 DR SMART: SM00410; Ig\_Like; 2.  
 DR PROSITE: PS00144; ASN\_GLN\_ASE\_1; UNKNOWN\_1.  
 KW Immunoglobulin domain.  
 SQ SEQUENCE 2016 AA; 222109 MW; 64A8DE3B7BD0AB0 CRC64;  
 Query Match 17.0%; Score 101; DB 5; Length 2016;  
 Best Local Similarity 29.2%; Pred. No. 0.06;  
 Matches 33; Conservative 18; Mismatches 48; Indels 14; Gaps 5;  
 QY 2 ONITARIGEPVLVKCGAPKPPORLEWKLNTGRTAEWKVLSPOGGGPPWSVARVLPNG- 60  
 Db 819 RNOTARGEPVAVLQCEAKGEPGIGL-WNNMNM-----LDPKNDNRYTIREEILSTGV 871  
 QY 61 -SLPLPAVGIDEGIPRCQAMNNGKETSRYRVYQIPGKP---EIVDSA 108  
 Db 872 MSSLSIKRTSRSDSALETCTVATNAFGSDAS-INMIVQEVPEMPYALKVLVDKS 923  
 RESULT 7  
 ID Q9BTNO PRELIMINARY; PRT; 628 AA.  
 AC Q9BTNO;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Hypothetical 66.3 kDa protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg K.;
RL Submitted (FEB-2001) to the EMBL/GenBank/Tran database
DR EMBL: BC004578; AAC04578.1;
DR InterPro: IPR004961; FN 111;
DR InterPro: IPR004599; Iq;
DR InterPro: IPR004598; Iq_c2;
DR InterPro: IPR004006; Iq_MHC;
DR InterPro: IPR001611; LRR;
DR InterPro: IPR000483; LRR_Cterm;
DR InterPro: IPR000472; LRR_Nterm;
DR InterPro: IPR003591; LRR_TYP;
DR Pfam: PF00041; fn3; 1;
DR Pfam: PF00047; Iq; 1;
DR Pfam: PF00560; LRR; 6;
DR Pfam: PF01463; LRRCT; 1;
DR PRINTS: PR00019; LEURICHRPT;
DR SMART: SM00060; FN3; 1;
DR SMART: SM00409; Iq; 1;
DR SMART: SM00408; Iq_c2; 1;
DR SMART: SM00082; LRRCT; 1;
DR SMART: SM00013; LRRNT; 1;
DR SMART: SM00369; LRR_TYP; 6;
KW Hypothetical protein; Immunoglobulin domain;
SQ SEQUENCE 628 AA: 66259 MW: 6855A5056AHC5FC GRC64;

Query Match 16.6%; Score 99; DB 4; Length 628;
Best Local Similarity 41.5%; Pred. No. 0.024;
Matches 34; Conservative 12; Mismatches 34; Indels 28; Gaps 6;

QY 9 GEPVLVLRKGGAPKPPQRLKWKLNIGRTKAWKVLSPQ--GGIDWDSVARVLNGSLFLP 65
DB 310 GEPALRCK-AGVDFEERVW ----- VSPQRLGN -SSKRAFFNGTLELL 454
QY 66 AVGLQDEGIFRCQAMNKNKETSNNYRVVYQI--- PKPEIVDSAS 109
DB 355 VIEPDGGGIFCTIAANAAGFAAA -VELTVGPPPPQPLANSTIS 396

RESULT 8
QYX141
ID QYX141 PRELIMINARY; PRI: 1248 AA;
AC QYX141;
DI 01-Nov-1999 (Tremblrel. 12; Created)
DT 01-Nov-1999 (Tremblrel. 12; Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19; Last annotation update)
DE Neural cell adhesion molecule L1.
GN L1CAM.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecus.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Fackell U., Schroder J., Resstet R., Veske A., Gul A.;
RL "Spectrum and detection rate of L1CAM mutations in isolated and familial cases with clinically suspected L1-Disease.";
DR EMBL: AF129167; AA028610.1;
DR HSSP: F20241; 1CFB;
DR InterPro: IPR004961; FN 111;
DR InterPro: IPR003598; Iq_c2;
DR InterPro: IPR004006; Iq_Like;
DR Pfam: PF00041; fn3; 5;
DR Pfam: PF00047; Iq; 6;
DR SMART: SM00060; FN3; 4;

Query Match 16.1%; Score 96; DB 6; Length 1018;
Best Local Similarity 26.4%; Pred. No. 0.091;
Matches 29; Conservative 26; Mismatches 35; Indels 20; Gaps 6;

QY 2 GNITAFLEPHVLRKQKAPKPPQRLKWKLNIGRTKAWKVLSPQ--GGIDWDSVARVLNGSLFLP 61
DB 249 KQVYALMGNVTRTP-FALGNVDPHRLWP -KVLQDM -DSTAFISTSCA 294
QY 62 LFLAVGLQDGGIFRCQAMNKNKETSNNYRVVYQIHWKPEIVDSASE 110
DB 244 VIKTINIGLFDGGTYPTFAENNP-RY-- RFGARY VQAPEWVEHIND 439

RESULT 10
QYX214
ID QYX214 PRELIMINARY; PRI: 1444 AA;
AC QYX214;

```



Fri May 30 17:16:34 2003

```

DR InterPro: IPR003600; Ig like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR000034; Laminin_B.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR002172; LDL_recept_A.
DR InterPro: IPR001993; Mitoch_carrier.
DR InterPro: IPR000519; P_trefoil.
DR Pfam: PF00047; iq; 16.
DR Pfam: PF00052; laminin_B; 2.
DR Pfam: PF00053; laminin_EGF; 5.
DR Pfam: PF00054; laminin_G; 2.
DR Pfam: PF00057; ldl_recept_a; 3.
DR PRINTS: PR00261; LDLRECEPTOR.
DR ProDom: PD003031; Laminin_B; 2.
DR SMART: SM00180; EGF_Lam; 5.
DR SMART: SM00001; EGF_Like; 3.
DR SMART: SM00408; IGC2; 12.
DR SMART: SM00410; IG_Like; 5.
DR SMART: SM00281; Lamb; 2.
DR SMART: SM00282; Lamb; 3.
DR SMART: SM00192; LDLa; 3.
DR SMART: SM00018; PD; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_4.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 7.
DR PROSITE: PS01209; LDLRA_1; 3.
DR PROSITE: PS00668; LDLRA_2; 4.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Alternative splicing; EGF-like domain; Glycoprotein;
KW Hypothetical protein; Immunoglobulin domain; Laminin EGF like domain;
KW Repeat.
KW REPEAT: 2482 AA; 270944 MW; R24006F7D6-94D70 CR664;
SQ
Query Match 16.1%; Score 95.5; DR 5; Length 2482;
Best local Similarity 25.48; Pred. No. 0.31;
Matches 31; Conservative 21; Mismatches 41; Indels 29; Gaps 5;
QY 8 IGEPVLKCK-GAPKKPPQRLKLTGRTEAMKVLSPQSG PWDSVAVLPNGSLFL 64
DB 1990 VGRPLQVKCAFCAQDPPEVFWLHDPC-----PERGDLDDPKPVTISEQPIRH 2040
QY 65 PAVGIQDEGIFRCQAMNNGKTKSNV-----KRVVYQIPG KPEIVDS 107
DB 2041 PNVGLNAGVYTCCKSSAIAIATKNIYFVVVFESKIATVGIAGNSQWFDGKEGELICT 2100
QY 108 AS 109
DB 2101 AT 2102
RESULT 13
Q9XT15 PRELIMINARY; PRI: 3375 AA.
AC Q9XT15;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE UNC-52 protein.
GN UNC-52 OR ZC101.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2.
RA Percy C., Barnes C.;
RL Submitted (MAP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS: FOUR FORMS: ISOFORM ZC101.2A (Q9XT15),
CC ISOFORM ZC101.2B (Q9B2G), ISOFORM ZC101.2C (Q9XT02) AND ISOFORM
CC ISOFORM ZC101.2E (SHOWN HERE); MAY BE PRODUCED BY ALTERNATIVE SPLICING.
DR EMBL: Z93395; CAB07708.1;
DR EMBL: Z93375; CAB07708.1; JOINED.
DR EMBL: Z93375; CAB07569.1;
DR EMBL: Z93395; CAB07569.1; JOINED.
DR HSP: P01130; I1DR.
DR WormPep, ZC101.2E; CE18424.

```

```

DR InterPro: IPR000561; EGF like.
DR InterPro: IPR003598; iq_c2.
DR InterPro: IPR003600; iq_like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR000034; Laminin_B.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR001993; Mitoch_carrier.
DR InterPro: IPR000519; P_trefoil.
DR Pfam: PF00047; iq; 17.
DR Pfam: PF00052; laminin_B; 2.
DR Pfam: PF00053; laminin_EGF; 5.
DR Pfam: PF00054; laminin_G; 2.
DR Pfam: PF00057; ldl_recept_a; 3.
DR PRINTS: PR00261; LDLRECEPTOR.
DR ProDom: PD003031; Laminin_B; 2.
DR SMART: SM00180; EGF_Lam; 5.
DR SMART: SM00001; EGF_Like; 3.
DR SMART: SM00408; IGC2; 12.
DR SMART: SM00410; IG_Like; 5.
DR SMART: SM00281; Lamb; 2.
DR SMART: SM00282; Lamb; 3.
DR SMART: SM00192; LDLa; 3.
DR SMART: SM00018; PD; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_7.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01248; LAMININ_TYPE_EGF; 7.
DR PROSITE: PS01209; LDLRA_1; 3.
DR PROSITE: PS00668; LDLRA_2; 4.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
KW Alternative splicing; EGF-like domain; Glycoprotein;
KW Hypothetical protein; Immunoglobulin domain; Laminin EGF like domain;
KW Repeat.
KW REPEAT: 3375 AA; 369049 MW; IAA418B14E5D67AA CR664;
SQ
Query Match 16.1%; Score 95.5; DR 5; Length 3375;
Best local Similarity 25.48; Pred. No. 0.45;
Matches 31; Conservative 21; Mismatches 41; Indels 29; Gaps 5;
QY 8 IGEPVLKCK-GAPKKPPQRLKLTGRTEAMKVLSPQSG PWDSVAVLPNGSLFL 64
DB 1990 VGRPLQVKCAFCAQDPPEVFWLHDPC-----PERGDLDDPKPVTISEQPIRH 2040
QY 65 PAVGIQDEGIFRCQAMNNGKTKSNV-----KRVVYQIPG KPEIVDS 107
DB 2041 PNVGLNAGVYTCCKSSAIAIATKNIYFVVVFESKIATVGIAGNSQWFDGKEGELICT 2100
QY 108 AS 109
DB 2101 AT 2102
RESULT 14
Q15402 PRELIMINARY; PRI: 182 AA.
AC Q15402;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE pregnancy-specific protein (fragment).
GN PSC1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS.
RX MEDLINE-96147764; PubMed-2392228;
RX Perjadin J., Tease J.A., Barnes W., Chan W.Y.;
RX "Expression of the pregnancy-specific beta 1-glycoprotein genes in
RX human testis.";

```

```

PI Biochem Biophys Res Commun 166:622-629(1990).
DR EMBL: M306729; AAA67524.1; -
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003600; Ig_like.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig; 2.
DR SMART: SM00408; IgC2; 1.
DR SMART: SM00410; Ig_Like; 1.
KW immunoglobulin domain.
FT NON_TER
SQ SEQUENCE 182 AA; 20274 MW; AC5663ADA412C5B0 CRC64;

Query Match: 16.0%; Score 95; DB 4; Length 182;
Best Local Similarity 27.8%; Pred. No. 0.614;
Matches 25; Conservative 12; Mismatches 31; Indels 22; Gaps 4.

QY 7 RIGPIVLKCKGAPKKPPQRLFWKINTGRTFAWKVLISGGGPPWISVAPVLPNGSLFLPA 66
DB 108 RSEVLYLST-SNDSNPPALYSWTIN-GKEQ-----LPGUKLFIPQ 146
QY 67 VGIQDEGIFHCQAMN-PNGKETKSNVYRVV 95
DB 147 ITTKHGLYACSVRSNATGKSSKSMTKV 176

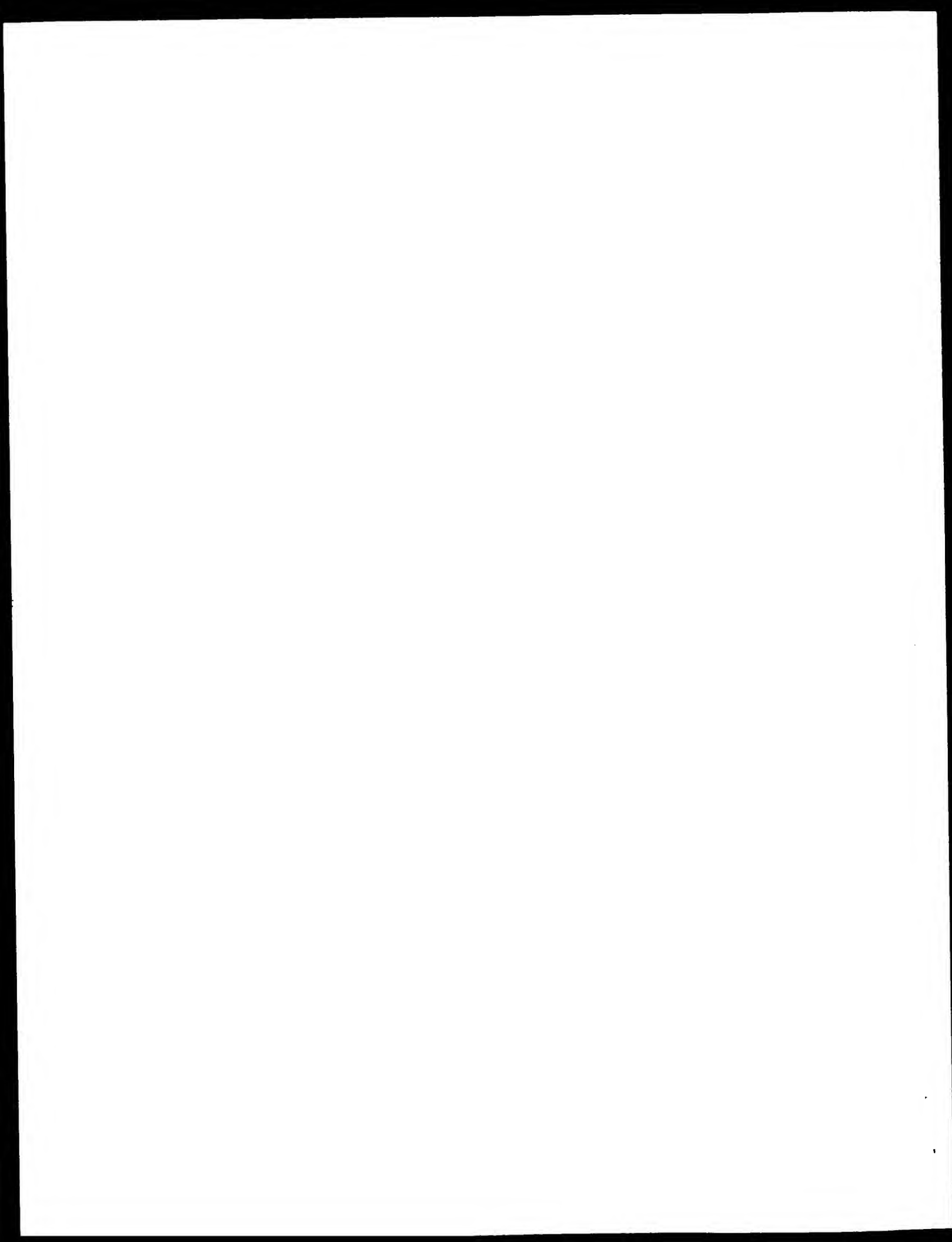
RESULT 15
Q96MS0
ID Q96MS0 PRELIMINARY: DPT: 1034 AA.
AC Q96MS0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CDNA FLJ11982 file, clone N128p7008550, highly similar to Mus musculus
DE rig-1 protein mRNA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Nimmiya K., Wajatsuna M., Kanda K., Kondo H., Yokoi I., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Futsu N., Sato K., Tanikawa M., Yamazaki M., Sugiura T., Ito R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Miyakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahara K., Masuho Y., Nagai K., Isegai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (Oct-2001) to the EMBL/GenBank/CCRC databases.
DR EMBL: AK056544; BA071212.1; -
DR InterPro: IPR003961; FN_111.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00041; fn3; 3.
DR Pfam: PF00047; Ig; 5.
SQ SEQUENCE 1034 AA; 110723 MW; 09E13C7B424F7E50 CRC64;

Query Match: 16.0%; Score 95; DB 4; Length 1034;
Best Local Similarity 27.9%; Pred. No. 0.12;
Matches 31; Conservative 11; Mismatches 35; Indels 34; Gaps 5;

QY 9 GEPLVLKCKGAPKKPPQRLFWKINTGRTFAWKVLISGGGPPWISVAPVLPNGSLFLPA 61
DB 78 GFPAFLPCR-AEGPPPINFWYKNGAPVATVREDPPAHRLI-----LPSGA 122
QY 62 LELPAV-----GIQDFGIFPQAMNPNGKETKSNVYRVV-----YQIPG 100
DB 123 LFFPRVHGRRAPDGEVWTGVAPNVLGAASPNASLEVAVLRLDDFRQSPG 173

Search completed: May 30, 2003, 15:57:28
Job time : 28.982 secs

```



OM protein - protein search, using sw model

Run on: May 30, 2003, 14:53:06 : Search time 75.5225 seconds  
(without alignments)  
585,775 Million cell updates/sec

Title: US-09-872-185b-2

Perfect score: 1766

Sequence: 1 AUNTAPR:EFVLEKGRAP

Scoring table: BIOSUM62

Gapop 10 0 0 Gapext 0 5

Searched: 508470 seqs, 18456620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A\_Geneseq\_101002.\*

- 1: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 2: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 3: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 4: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 5: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 6: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 7: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 9: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 10: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 11: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 12: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 13: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 14: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 15: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 16: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 17: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 18: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 19: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 20: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 21: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 22: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 23: /SIDS2/qqdata/geneseq/geneseq-emb1/AA1986.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1766	100.0	404	23	AAE23218
2	1766	100.0	404	23	AAU77542
3	1762	99.8	404	23	AAU77544
4	1762	99.8	404	23	AAU77543
5	1762	99.8	404	23	AAU77543
6	1701	96.3	318	18	AAW3754
7	1701	96.3	318	18	AAW3754
8	1701	96.3	318	18	AAW3754
9	1701	96.3	318	18	AAW3754
10	1696	96.0	339	23	AAU48746

11	1463	82.8	416	23	AAE23218	cow receptor for a
12	1463	82.8	416	23	AAU77542	bovine receptor for
13	1383.5	78.3	403	23	AAU77544	mouse receptor for
14	1383.5	78.3	403	23	AAU77544	mouse receptor for
15	595	33.7	112	23	AAU48747	human RAGE protein
16	220.5	12.5	582	22	AAU76858	human lung tumour
17	220.5	12.5	582	22	AAU76858	human lung tumour
18	218.5	12.4	583	17	AAU06891	clone #18964 of lu
19	218.5	12.4	583	19	AAU47088	human activated leucocyte
20	217.5	12.3	646	20	AAU42405	human MUC18 amino
21	217.5	12.3	646	20	AAU42405	human MUC18 amino
22	216.5	12.3	646	20	AAU42405	human MUC18 amino
23	214.5	12.1	570	17	AAU77542	stem cell marker H
24	214.5	12.1	570	17	AAU77542	stem cell marker H
25	212.5	12.0	753	20	AAU83927	human T85 protein.
26	210	11.9	1496	20	AAU81030	melanoma associate
27	210	11.9	1496	20	AAU70459	human p53 target m
28	210	11.9	1496	20	AAU70459	human p53 target m
29	209	11.8	1449	20	AAU04404	human peroxidase
30	209	11.8	1449	20	AAU04404	human peroxidase
31	208.5	11.8	1449	20	AAU04404	human peroxidase
32	204.5	11.6	511	16	AAU77438	human pobo 1 polyp
33	204.5	11.6	511	16	AAU77438	human pobo 1 polyp
34	203	11.5	1033	22	ABU64123	hgp (1-314)/CEA (4
35	200.5	11.4	509	16	AAU77437	novel human protei
36	200.5	11.4	509	16	AAU77437	novel human protei
37	197.5	11.2	404	22	AAU61142	protophila melanoq
38	196.5	11.1	206	23	ABU42182	hgp (1-314)/CEA (4
39	196.5	11.1	782	23	ABU53278	human ovarian anti
40	196.5	11.1	844	23	ABU53277	human polypeptide
41	194.5	11.0	898	22	ABU12152	human polypeptide
42	194	10.9	1483	22	ABU12152	novel human diagno
43	191.5	10.8	628	9	AAU81229	novel human diagno
44	190.5	10.8	679	23	AAU39822	carcinoembryonic a
45	190.5	10.8	698	16	AAU65168	carcinoembryonic a

## ALIGNMENTS

RESULT 1  
AAU61925  
ID: AAU61925 standard; protein; 404 AA.  
AC: AAU61925;  
XX: AAU61925;  
DE: 15 JUN 2001 (first entry)  
XX: Extracorporeal circulation material receptor protein.  
XX: Extracorporeal circulation; carbonyl stress product; receptor;  
KW: diabetes; vascular lesion; excretory dysfunction.  
XX: Unidentified.  
OS: Unidentified.  
PN: W0200118060-AA.  
XX: 15-MAR-2001.  
PF: 08-SEP-2000; 2000CW-3P06172.  
PP: 08-SEP-1999; 99JP-0254463.  
XX: (CDBA) TORAY INC.  
XX: Shimizu S, Kikuta M, Akiyama H, Usui M.  
DR: WPI: 2001-290314/30.  
XX: Material for extracorporeal circulation, applicable in selective  
PI: elimination of diabetic complication factors such as carbonyl stress  
PT: products caused by abnormally promoted carbonyl stress from excretory  
PT: dysfunction in vascular lesions.

Material for extracorporeal circulation, applicable in selective  
elimination of diabetic complication factors such as carbonyl stress  
products caused by abnormally promoted carbonyl stress from excretory  
dysfunction in vascular lesions.

Claim 1: Page 31 42: 46pp: Japanese.

The present invention describes a material for extracorporeal circulation which is made from a water insoluble carrier immobilized with a protein having the sequence shown here. The materials of the invention, including adsorbents, are for extracorporeal circulation, which are applicable in the selective elimination of diabetic complication factors from a body fluid, and are therefore useful in treating vascular lesions like arteriosclerosis due to carbonyl stress products caused by abnormally increased carbonyl stress from excretory dysfunction.

XX	Sequence	404 AA:	Query Match	100.0%	Score 1766:	DB 22:	length 404:
SQ			Best Local Similarity	100.0%	Pred. No. 1,26-124:		
	Matches 332:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:		
QY	1	AQNTIAPIGPELVIKKGKAPFPQPPQPLEFKLNTGPTFAWKVLSPVGGGQPMISVARVL	PNG	60			
DB	23	AQNTIAPIGPELVIKKGKAPFPQPPQPLEFKLNTGPTFAWKVLSPVGGGQPMISVARVL	PNG	82			
QY	61	SLFLPVGIGDEGIFPCQAMNNGKETKSNYRVRYQIPGKPEIVDSASELTACVPNKVG	120				
DB	83	SLFLPVGIGDEGIFPCQAMNNGKETKSNYRVRYQIPGKPEIVDSASELTACVPNKVG	142				
QY	121	TCVSEGSYPACTLSWHLLCKPLVPNKGVSVKQETRRHPEITCLPTQSELMVTPARGGDP	180				
DB	143	TCVSEGSYPACTLSWHLLCKPLVPNKGVSVKQETRRHPEITCLPTQSELMVTPARGGDP	202				
QY	181	KPTSCSFSGLPFHRIALRTAPITQPRVWPEPVEEVQLVWPEEGVAVAPGGTIVTLTCEVP	240				
DB	203	KPTSCSFSGLPFHRIALRTAPITQPRVWPEPVEEVQLVWPEEGVAVAPGGTIVTLTCEVP	262				
QY	241	AQPSQPIHWKKDGVPLPLPSPVLLIPEIQPODQCTYSCVATFSSHGQPEKRAVYSISIE	300				
DB	263	AQPSQPIHWKKDGVPLPLPSPVLLIPEIQPODQCTYSCVATFSSHGQPEKRAVYSISIE	322				
QY	301	PGREGPTAGSVGGSGIATLALGILGGIGTA	312				
DB	323	PGREGPTAGSVGGSGIATLALGILGGIGTA	354				

RESULT 2	
AA048745	
ID	AA048745 standard; protein: 404 AA.
XX	
AC	AA048745;
XX	
DT	02-APR-2002 (first entry)
XX	
DE	Human RAGE protein SEQ ID NO 1.
XX	
KW	human; RAGE; receptor for advanced glycosylated endproduct; receptor;
KW	antidiabetic; neuroprotective; cytostatic; antiinflammatory; vasotropic;
KW	nephropathic; dermatological; antihypertensive; antidiabetic; diabetes;
KW	Alzheimer's disease; cancer; inflammation; kidney failure;
KW	systemic lupus; nephritis; erectile dysfunction; atherosclerosis.
XX	
OS	Homo sapiens.
XX	
FN	W0200192892-A2.
XX	
PD	06-DEC-2001.
XX	
PF	30-MAY-2001; 2001WO-0517447.
XX	
PR	30-MAY-2000; 2000US-207342P.
PR	05-MAR-2001; 2001US-0799152.
XX	
PA	(TRAN-) TRANS TECH PHARMA.
PA	Shabbaz M;
P1	

XX	WP1: 2982 1:4374/15.
XX	Detecting a receptor for advanced glycosylated endproducts (RAGE)
XX	modulators, for treating e.g., cancer, diabetes or inflammation,
PPT	comprises measuring the amount of bound anti RAGE antibody
XX	Claim 1; Fig 2; 4ppp; English.
CC	The invention relates to detecting receptor for advanced glycosylated
CC	endproducts (RAGE) modulators comprises determining the amount of RAGE
CC	protein or its fragment bound to the pre-adsorbed liquid by measuring the
CC	amount of anti-RAGE antibody bound to the solid surface. The method is
CC	useful for rapid, high-throughput identification of compounds that
CC	modulate RAGE. The compounds are useful for treating symptoms of diabetes
CC	and symptoms of diabetic late complications, amyloidosis, Alzheimer's
CC	disease, cancer, inflammation, kidney failure, systemic lupus nephritis
CC	or inflammatory lupus nephritis, erectile dysfunction and
CC	atherosclerosis.
XX	Sequence 404 AA;
QC	Query Match 100.0%; Score 1766; DB 24; Length 404;
	Best Local Similarity 150.0%; Fred. No. 1,2e+124;
Matches 432; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 AGNITARIGELHLVKIKGAPKKPPQLEKKNLNGRIAMKVLSGGGIIWISVAHVLFG 60
DB	
QY	23 AGNITARIGELHLVLRCKCAIKPKPQLEKKNIGRTFAKKVLSPOGCGIWISSVAHVLFG 82
DB	
QY	61 SLFLPAVGILQEGTFRQQANRNCKETKSNRYRVVQIPGRKEIVISASETATGVPNKVG 120
DB	
QY	83 SLFLPAVGILQEGTFRQQANRNCKETKSNRYRVVQIPGRKEIVISASETATGVPNKVG 142
DB	
QY	121 TCVSCHSPACTLSWHLLDPLFVNPKGSVKRGTHRRPELLEFTLDSELMVTARGSDP 180
DB	
CL	143 TCVESEGSPATLSWHLLDPLFVNPKGSVKRGTPPHPETQLFTLDSELMVTARGSDP 202
QY	181 RPTSCSPSGLPRRHALPTAPIRWVPVPIERVQI VVEDPGCAVAGCTIVTLFCRP 240
DB	
QY	203 RPTSCSPSGLPRRHALPTAPIRWVPVPIERVQI VVEDPGCAVAGCTIVTLFCRP 262
DB	
QY	241 AQPSPQHMMKGDPGLPLPSPVLIIPETGPQGCTYSCTVAHSISBGPESHAVSISILE 400
DB	
QY	263 AQPSPQHMMKGDPGLPLPSPVLIIPETGPQGCTYSCTVA HSSHGTGPESHAVSISILE 422
DB	
QY	301 PGEPTACSGVSGSLGTLALALGLAGLATA 432
DB	
QY	423 PGEPTACSGVSGSLGTLALALGLAGLATA 454
DB	
RESULT 4	
ID	AA552140
XX	AA552140 standard; protein; 432 AA.
XX	AA552140;
XX	28-JAN-2000 (first entry)
UT	Human Receptor to AGE (RAGE) amino acid sequence.
DE	Soluble receptor for advanced glycation endproducts; RAGE; tumour;
XX	invasion; metastasis; amphoterin; neuron; inhibit therapy.
KW	
WM	

RESULT 4	
AA52130	
ID	AA52130 standard; protein; 332 AA.
XX	
AC	AA52130;
XX	
DT	28-JAN-2000 (first entry)
XX	
DE	Human Receptor to AGE (RAGE) amino acid sequence.
XX	
KW	Soluble receptor for advanced glycation endproducts; RAGE; tumour;
KW	invasion; metastasis; chemokine; neuron; inhibit. therapy.



XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 PA Schmidt AM, Stern D;  
 PI WPI: 2868 61426961  
 XX Inhibiting tumour invasion or spreading by administration of soluble  
 PT receptor for advanced glycation endproducts  
 XX  
 PS Disclosure: Page 10-11; 88pp; English.  
 XX  
 CC This is the amino acid sequence of the human soluble Receptor for  
 CC Advanced Glycation Endproducts (RAGE). RAGE interacts with a range of  
 CC physiologically and pathophysiologically relevant ligands when  
 CC considering tumour invasion in normal developing neurons. RAGE  
 CC colocalizes with amyloid in which is a matrix associated polypeptide.  
 CC The expression of both RAGE and amyloid increases after birth, but  
 CC both have increased expression in tumours. RAGE polypeptides  
 CC AY52132-52135 are used in the invention in a method for inhibiting  
 CC tumour invasion and metastasis. The method involves inhibiting tumour  
 CC invasion and metastasis via administration of a therapeutically effective  
 CC amount of the pharmaceutical composition containing a RAGE polypeptide.  
 CC The invention also relates to a method for evaluating the ability of an  
 CC agent to inhibit tumour invasion in a local cellular environment. RAGE  
 CC can be administered to a patient in a pharmaceutically acceptable  
 CC carrier.  
 XX  
 SQ Sequence 332 AA;  
 Query Match: 99.8%, Score 1763, DB 21, Length 332,  
 Best Local Similarity 99.7%, Pred. No. 1.5e-124;  
 Matches 331; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AQNITARIGEPVLVKKGAPKPPQRLKLNKTCRTKAWKVLSPQGGGWDVSVARVLPNG 60  
 DB 1 AQNITARIGEPVLVKKGAPKPPQRLKLNKTCRTKAWKVLSPQGGGWDVSVARVLPNG 60  
 QY 61 SLFLPAVGIDQEGIFRCQAMNRNCKETKSNRYRVVYQIPCKPEIIVDSASSETAGVFNKVG 120  
 DB 61 SLFLPAVGIDQEGIFRCQAMNRNCKETKSNRYRVVYQIPCKPEIIVDSASSETAGVFNKVG 120  
 QY 121 TCVSEGSYPAGTILSWHLDCKPLVFNKGVSKQETRRHPTGLTQLQSELMTVPARGGDP 180  
 DB 121 TCVSEGSYPAGTILSWHLDCKPLVFNKGVSKQETRRHPTGLTQLQSELMTVPARGGDP 180  
 QY 181 RPTESCSPGLPRLPALPTAPQPRVWEPVPLEEVQLVVEPEGCAVAPGCTVTLTCEVP 240  
 DB 181 RPTESCSPGLPRLPALPTAPQPRVWEPVPLEEVQLVVEPEGCAVAPGCTVTLTCEVP 240  
 QY 241 AQPSTQIHWMMKDGVPPLPSPVLLIPETGPDQGTGYSVATHSSHGPOESRAVSISTIE 300  
 DB 241 AQPSTQIHWMMKDGVPPLPSPVLLIPETGPDQGTGYSVATHSSHGPOESRAVSISTIE 300  
 QY 301 PCEGPTAGSGVSGGLCTIALALGILGGLGTA 332  
 DB 301 PCEGPTAGSGVSGGLCTIALALGILGGLGTA 332  
 RESULT 4  
 ID AAE23219 standard; Protein: 404 AA.  
 XX  
 AC AAE23219;  
 XX  
 DT 27-AUG-2002 (first entry)  
 DE Human receptor for advanced glycosylation end product (PAGE) protein.  
 KW Human, Receptor for advanced glycosylation end product; RAGE; cardiant;  
 KW tissue growth, neointimal formation, blood vessel; restenosis; diabetes;  
 KW myocardial infarction; aneuploidy; peripheral vascular surgery; angina;  
 KW transgenic animal; acute thrombotic stroke; venous thrombosis.

XX OS Homo sapiens.  
 XX WO200230889-A2.  
 XX 18-APR-2002.  
 XX 12 OCT 2001; 2001WO-US-0036.  
 XX 13-OCT-2000; 2000US-0687528.  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 PA Stern DM, Schmidt A, Marso S, Topol E, Lincoff AM;  
 PI WPI: 2002 425260/45.  
 DE N-PSDB: AAU30952.  
 XX Inhibiting new tissue growth or neointimal formation in blood vessels  
 PI of subject suffering from diabetes, stroke and preventing restenosis,  
 PI comprises administering inhibitor of receptor for advance glycation end  
 XX product  
 ES Disclosure, Page 16; 43pp; English.  
 XX The invention relates to a method for inhibiting new tissue growth or  
 CC neointimal formation in blood vessels in a subject that has experienced  
 CC blood vessel injury and preventing exaggerated restenosis in a diabetic  
 CC subject. The method comprises administering an inhibitor of receptor for  
 CC advanced glycation/glycosylation end product (RAGE), so as to inhibit new  
 CC tissue growth or neointimal formation in subject's blood vessels and  
 CC preventing restenosis in the subject. The method is useful for inhibiting  
 CC new tissue growth or neointimal formation in blood vessels in a subject  
 CC like non-human animal, a transgenic non-human animal or a human suffering  
 CC from diabetes, acute thrombotic stroke, venous thrombosis, unstable  
 CC angina, myocardial infarction, abrupt closure following angioplasty or  
 CC stent placement, or thrombosis as a result of peripheral vascular surgery.  
 CC The method is also useful for preventing restenosis and for determining  
 CC whether a compound inhibits new tissue growth in a blood vessel in a  
 CC subject. The present sequence is human receptor for advanced  
 CC glycosylation end product (RAGE) protein.  
 XX  
 SQ Sequence 404 AA;  
 Query Match: 99.8%, Score 1762, DB 23, Length 404;  
 Best Local Similarity 99.7%, Pred. No. 2.3e-124;  
 Matches 331; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AQNITARIGEPVLVKKGAPKPPQRLKLNKTCRTKAWKVLSPQGGGWDVSVARVLPNG 60  
 DB 23 AQNITARIGEPVLVKKGAPKPPQRLKLNKTCRTKAWKVLSPQGGGWDVSVARVLPNG 82  
 QY 61 SLFLPAVGIDQEGIFRCQAMNRNCKETKSNRYRVVYQIPCKPEIIVDSASSETAGVFNKVG 120  
 DB 83 SLFLPAVGIDQEGIFRCQAMNRNCKETKSNRYRVVYQIPCKPEIIVDSASSETAGVFNKVG 142  
 QY 121 TCVSEGSYPAGTILSWHLDCKPLVFNKGVSKQETRRHPTGLTQLQSELMTVPARGGDP 180  
 DB 143 TCVSEGSYPAGTILSWHLDCKPLVFNKGVSKQETRRHPTGLTQLQSELMTVPARGGDP 202  
 QY 181 RPTESCSPGLPRLPALPTAPQPRVWEPVPLEEVQLVVEPEGCAVAPGCTVTLTCEVP 240  
 DB 203 RPTESCSPGLPRLPALPTAPQPRVWEPVPLEEVQLVVEPEGCAVAPGCTVTLTCEVP 262  
 QY 241 AQPSTQIHWMMKDGVPPLPSPVLLIPETGPDQGTGYSVATHSSHGPOESRAVSISTIE 300  
 DB 263 AQPSTQIHWMMKDGVPPLPSPVLLIPETGPDQGTGYSVATHSSHGPOESRAVSISTIE 322  
 QY 301 PCEGPTAGSGVSGGLCTIALALGILGGLGTA 332  
 DB 323 PCEGPTAGSGVSGGLCTIALALGILGGLGTA 354



QY 1 AQTITARIGETVLKTKAPKPPQPLEWKLNTPTAEWKVLSPGSGPWSVAVPLNG 60  
 DB 1 AQTITARIGETVLKTKAPKPPQPLEWKLNTPTAEWKVLSPGSGPWSVAVPLNG 60  
 QY 61 SLFLPAVGTQDEIFKQAMNPNKTKSNRYRVVQIPKPEIVDSASELTAGVFNKG 120  
 DB 61 SLFLPAVGTQDEIFKQAMNPNKTKSNRYRVVQIPKPEIVDSASELTAGVFNKG 120  
 QY 121 TCVEGSPAGTILSWHLDGKPLVNEKGVSVKQPKRHPETGLFTLOSELMTVPARGGDP 180  
 DB 121 TCVEGSPAGTILSWHLDGKPLVNEKGVSVKQPKRHPETGLFTLOSELMTVPARGGDP 180  
 QY 181 PPTFSCSFSPGLPHHPALPTAPQPPWEPVPLEEVQIVVEPEGJAVAPSGTITLTCEVP 240  
 DB 181 PPTFSCSFSPGLPHHPALPTAPQPPWEPVPLEEVQIVVEPEGJAVAPSGTITLTCEVP 240  
 QY 241 AQSPQTHWKKGVPLPLPSPVLLPEIGPQGGIYSCVATHSSHGPOESRAVSISIE 300  
 DB 241 AQSPQTHWKKGVPLPLPSPVLLPEIGPQGGIYSCVATHSSHGPOESRAVSISIE 300  
 QY 301 PGEESPTAGSVGGSGIGT 318  
 DB 301 PGEESPTAGSVGGSGIGT 318

## RESULT 7

AAW33754  
 ID AAW33754 standard; protein: 318 AA.

XX AAW33754;

AC 08-MAY-1998 (first entry)

DE Human RAGE polypeptide (318 amino acid residues).

XX Advanced glycosylation end product receptor, RAGE, screening; AGE;  
 KW vascular permeability, diabetes mellitus, treatment; atherosclerosis;  
 KW Alzheimer's disease.

OS Homo sapiens.

XX W09739121-AL.

XX 23 OCT-1997.

XX 11-APR-1997; 97WO-EP01842.

XX 16-APR-1996; 96US-0633147.

XX (SCHD) SCHERING AG.

XX Morser MJ, Nagashima M;

XX WPI; 1997-526458/48.

XX N-PSDB; AAV06518.

XX New soluble advanced glycosylation end-product receptor polypeptide  
 PT - used for reducing vascular permeability, complications of diabetes  
 PT etc., also for purification and to screen for modulators

XX Claim 3; Fig 1B; 91pp; English.

XX This is a human advanced glycosylation end product receptor (RAGE)  
 CC polypeptide (318 amino acid residues). The RAGE polypeptides and its  
 CC active fragments or their mimetics, inhibit interaction between advanced  
 CC glycosylation end-products (AGE) and a receptor (specifically RAGE). They  
 CC are used to treat diseases associated with AGE/RAGE interaction, such as  
 CC increased vascular permeability, diabetes mellitus (particularly  
 CC complications such as micro- or macro-vasculopathy or occlusive vascular  
 CC disorders such as neuropathy, nephropathy, retinopathy or  
 CC atherosclerosis) or hemodialysis-associated amyloidosis, also activation  
 CC of microglial cells by beta-amyloid peptides in Alzheimer's disease or

CC age-related disorders such as oxidative stress. These RAGE polypeptides  
 CC are also used, when immobilised, to purify AGE from a protein mixture and  
 CC to screen for compounds that are agonists and antagonists of AGE/RAGE  
 CC interaction. They can also be used diagnostically to detect abnormal  
 CC levels of AGE. Antibodies against RAGE polypeptides are useful as  
 CC immunoassay reagents for measurement of RAGE levels, and as inhibitors of  
 CC interaction between AGE and RAGE or other receptors and for purification  
 CC and quantitation of RAGE polypeptides. The encoding nucleic acids are  
 CC used to express recombinant RAGE and as probes for isolating related  
 CC genes.

XX Sequence 318 AA;

Query Match 95.3%; Score 1781, DB 18, Length 318;

Best local similarity 100.0%; Pred. No. 6.7e-120;

Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AQTITARIGETVLKTKAPKPPQPLEWKLNTPTAEWKVLSPGSGPWSVAVPLNG 60

DB 1 AQTITARIGETVLKTKAPKPPQPLEWKLNTPTAEWKVLSPGSGPWSVAVPLNG 60

QY 61 SLFLPAVGTQDEIFKQAMNPNKTKSNRYRVVQIPKPEIVDSASELTAGVFNKG 120

DB 61 SLFLPAVGTQDEIFKQAMNPNKTKSNRYRVVQIPKPEIVDSASELTAGVFNKG 120

QY 121 TCVEGSPAGTILSWHLDGKPLVNEKGVSVKQPKRHPETGLFTLOSELMTVPARGGDP 180

DB 121 TCVEGSPAGTILSWHLDGKPLVNEKGVSVKQPKRHPETGLFTLOSELMTVPARGGDP 180

QY 181 PPTFSCSFSPGLPHHPALPTAPQPPWEPVPLEEVQIVVEPEGJAVAPSGTITLTCEVP 240

DB 181 PPTFSCSFSPGLPHHPALPTAPQPPWEPVPLEEVQIVVEPEGJAVAPSGTITLTCEVP 240

QY 241 AQSPQTHWKKGVPLPLPSPVLLPEIGPQGGIYSCVATHSSHGPOESRAVSISIE 300

DB 241 AQSPQTHWKKGVPLPLPSPVLLPEIGPQGGIYSCVATHSSHGPOESRAVSISIE 300

QY 301 PGEESPTAGSVGGSGIGT 318

DB 301 PGEESPTAGSVGGSGIGT 318

## RESULT 8

AAW44199

ID AAW44199 standard; protein: 340 AA.

XX AAW44199;

AC 14-MAY-1998 (first entry)

DE Human soluble receptor to an advanced glycosylation end product.

XX Human, soluble receptor, advanced glycosylation end product, RAGE;

XX AGE; antibody; vascular permeability; diabetes mellitus.

XX Homo sapiens.

XX W09739125-AL.

XX 23-OCT-1997.

XX 11-APR-1997; 97WO-EP01834.

XX 16-APR-1996; 96US-0633148.

XX (SCHD) SCHERING PATENTE AG.

XX Hollander DA, Morser MJ, Nagashima M;

XX WPI; 1997-558580/51.

XX N-PSDB; AAV12394.

XX Anti-advanced glycosylation end product polypeptide antibody

PT prevents receptor binding and therefore reduces vascular  
 permeability, useful to treat diabetes mellitus  
 PS Claim 2: Page 40-41; 90pp; English.  
 XX The present sequence represents a soluble human receptor to an advanced  
 CC glycosylation end product (RAGE) polypeptide. The present invention  
 CC describes an isolated antibody (Ab), specifically immunoreactive with  
 CC RAGE. Advanced glycosylation end products (AGE) of proteins are  
 CC non-enzymatically glycosylated proteins, which accumulate in vascular  
 CC tissue in ageing, and at an accelerated rate in individuals with  
 CC diabetes. The Ab, which prevents the interaction between an AGE and it's  
 CC receptor (RAGE), reduces vascular permeability. The Ab can be used to  
 CC treat diabetes mellitus symptoms, e.g. microvasculopathy, occlusive  
 CC vascular disorders, neuropathy, nephropathy, retinopathy, haemodialysis  
 CC associated amyloidosis or atherosclerosis. The Ab can also be used for  
 CC the isolation and purification of human RAGE polypeptide.  
 XX Sequence: 440 AA;

Query Match 96.4%; Score 1701; DB 18; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 7, 36, 120;  
 Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AGNITARIGEDLVLRKQAPKKPPPLEWKLNTGRTFAWKVLSPGGGPMWSVAPVLPNG 60  
 DB 23 AGNITARIGEDLVLRKQAPKKPPPLEWKLNTGRTFAWKVLSPGGGPMWSVAPVLPNG 62  
 QY 61 SLFLPFAVGIDQEGIFPCVAMNPNCKETKSNYPVPVYGLPCKPDLVLSASSELTAQVKNVG 120  
 DB 83 SLFLPFAVGIDQEGIFPCVAMNPNCKETKSNYPVPVYGLPCKPDLVLSASSELTAQVKNVG 142  
 QY 121 TCVSFGSYFACHTASWHLDPKPLVPNPKGVSVVETTPPHEGLTQSELMVTPARGGDP 180  
 DB 143 TCVSFGSYFACHTASWHLDPKPLVPNPKGVSVVETTPPHEGLTQSELMVTPARGGDP 202  
 QY 181 RPTSCSFSPGLPRRALPTAPLQPVNPKGVSVVETTPPHEGLTQSELMVTPARGGDP 240  
 DB 204 RPTSCSFSPGLPRRALPTAPLQPVNPKGVSVVETTPPHEGLTQSELMVTPARGGDP 262  
 QY 241 AQPSTQTHMKKGVPLPPLPSVLLPPLGPGQGYTSVATHSSHPQPSRAVSTIE 300  
 DB 263 AQPSTQTHMKKGVPLPPLPSVLLPPLGPGQGYTSVATHSSHPQPSRAVSTIE 322  
 QY 401 PGEQPTAGSVGSGSLGT 418  
 DB 423 PGEQPTAGSVGSGSLGT 440

RESULT 9  
 AAW43754  
 XX AAW43754 standard; protein: 440 AA.  
 AC AAW43754;  
 XX  
 DT 08 MAY 1998 (first entry)  
 XX  
 DE Human RAGE polypeptide (440 amino acid residues).  
 XX  
 KW Advanced glycosylation end product receptor; RAGE; screening; AGE;  
 FW vascular permeability; diabetes mellitus; treatment; atherosclerosis;  
 XX Alzheimer's disease.  
 XX  
 OS Homo sapiens.  
 XX  
 EN W09739121-A1.  
 XX  
 PD 24 OCT 1997.  
 XX  
 PF 11 APR 1997; 97WO EP01842.  
 XX  
 PR 16 APR 1996; 96US 064147.  
 XX

FA (Scheink, A.,  
 XX Morser MJ, Nadeshima M;  
 XX WPL 1997 526458/48;  
 DB N-PSUB; AAV06517.  
 XX New soluble advanced glycosylation end product receptor polypeptide  
 PT used for reducing vascular permeability, complications of diabetes  
 PI etc., also for purification and to screen for modulators  
 XX Claim 3: Fig 1A; 91pp; English.

This is a human advanced glycosylation end product receptor (RAGE)  
 polypeptide (340 amino acid residues). The RAGE polypeptides and its  
 active fragments or their mimetics, inhibit interaction between advanced  
 glycosylation end products (AGE) and a receptor (specifically RAGE). They  
 are used to treat diseases associated with AGE/RAGE interaction, such as  
 increased vascular permeability, diabetes mellitus (particularly  
 complications such as micro- or macro-vasculopathy or occlusive vascular  
 disorders such as neuropathy, nephropathy, retinopathy or  
 atherosclerosis) or haemodialysis-associated amyloidosis, also activation  
 of microglial cells by beta amyloid peptides in Alzheimer's disease or  
 age-related disorders such as oxidative stress. These RAGE polypeptides  
 are also used, when immobilised, to purify AGE from a protein mixture and  
 to screen for compounds that are agonists and antagonists of AGE/RAGE  
 interaction. They can also be used diagnostically to detect abnormal  
 levels of AGE. Antibodies against RAGE polypeptides are useful as  
 immunoassay reagents for measurement of RAGE levels, and as inhibitors of  
 interaction between AGE and RAGE or other receptors, and for purification  
 and quantification of RAGE polypeptides. The encoding nucleic acids are  
 used to express recombinant RAGE and as probes for isolating related  
 genes.

Sequence: 440 AA;  
 Query Match 96.4%; Score 1701; DB 18; Length 440;  
 Best Local Similarity 100.0%; Pred. No. 7, 36, 120;  
 Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGNITARIGEDLVLRKQAPKKPPPLEWKLNTGRTFAWKVLSPGGGPMWSVAPVLPNG 60  
 DB 23 AGNITARIGEDLVLRKQAPKKPPPLEWKLNTGRTFAWKVLSPGGGPMWSVAPVLPNG 62  
 QY 61 SLFLPFAVGIDQEGIFPCVAMNPNCKETKSNYPVPVYGLPCKPDLVLSASSELTAQVKNVG 120  
 DB 83 SLFLPFAVGIDQEGIFPCVAMNPNCKETKSNYPVPVYGLPCKPDLVLSASSELTAQVKNVG 142  
 QY 121 TCVSFGSYFACHTASWHLDPKPLVPNPKGVSVVETTPPHEGLTQSELMVTPARGGDP 180  
 DB 143 TCVSFGSYFACHTASWHLDPKPLVPNPKGVSVVETTPPHEGLTQSELMVTPARGGDP 202  
 QY 181 RPTSCSFSPGLPRRALPTAPLQPVNPKGVSVVETTPPHEGLTQSELMVTPARGGDP 240  
 DB 204 RPTSCSFSPGLPRRALPTAPLQPVNPKGVSVVETTPPHEGLTQSELMVTPARGGDP 262  
 QY 241 AQPSTQTHMKKGVPLPPLPSVLLPPLGPGQGYTSVATHSSHPQPSRAVSTIE 300  
 DB 263 AQPSTQTHMKKGVPLPPLPSVLLPPLGPGQGYTSVATHSSHPQPSRAVSTIE 322  
 QY 401 PGEQPTAGSVGSGSLGT 418  
 DB 423 PGEQPTAGSVGSGSLGT 440

RESULT 10  
 AAM48746  
 ID AAM48746 standard; protein: 440 AA  
 XX  
 AC AAM48746;  
 XX  
 DT 02 APR 2002 (first entry)  
 XX

DE Human SPAGE protein SEQ ID NO 2.  
 XX  
 KW Human: RAGE; receptor for advanced glycosylated endproduct, receptor;  
 KW antidiabetic, neuroprotective, cytostatic; antiinflammatory; vasotrophic;  
 KW nephrotropic, dermatologic; antiatherosclerotic, neurotropic; diabetes;  
 KW Alzheimer's disease; cancer; inflammation; kidney failure;  
 KW systemic lupus; nephritis; erectile dysfunction; atherosclerosis.  
 OS  
 XX Homo sapiens.  
 XX  
 PN WO2001:92892-A2.  
 PD 06-DEC-2001  
 XX  
 XX 30-MAY-2001: 2001WO-US17447.  
 XX  
 PR 05-MAY-2000: 2000US-207342P.  
 PP 05-MAY-2001: 2001US-076912.  
 XX  
 PA (TRAN-) TRANS TECH PHARMA.  
 XX  
 PI Shahbaz M;  
 XX  
 DR WPI: 2002-114372/15.  
 XX  
 PT Detecting a receptor for advanced glycosylated endproducts (RAGE)  
 PT modulators, for treating e.g., cancer, diabetes or inflammation,  
 PT comprises measuring the amount of bound anti-RAGE antibody -  
 XX  
 PS Claim 2: Fig 2: 49pp: English.  
 XX  
 CC The invention relates to detecting receptor for advanced glycosylated  
 CC endproducts (RAGE) modulators comprising determining the amount of RAGE  
 CC protein or its fragment bound to the pre-adsorbed ligand by measuring the  
 CC amount of anti-RAGE antibody bound to the solid surface. The method is  
 CC useful for rapid, high-throughput identification of compounds that  
 CC modulate RAGE. The compounds are useful for treating symptoms of diabetes  
 CC and symptoms of diabetic late complications, amyloidosis, Alzheimer's  
 CC disease, cancer, inflammation, kidney failure, systemic lupus nephritis  
 CC or inflammatory lupus nephritis, erectile dysfunction and  
 CC atherosclerosis.  
 XX  
 XX Sequence 339 AA;  
 XX  
 Query Match 96.0%; Score 1696; DB 23; Length 339;  
 Best Local Similarity 100.0%; Pred No 1,7e-119;  
 Matches 317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 AQNITARIGEPVLKCKGAPKPPORLEWKLNTGRTAEAKVLSPOGGPMDSVARVLPNG 60  
 DB 23 AQNITARIGEPVLKCKGAPKPPORLEWKLNTGRTAEAKVLSPOGGPMDSVARVLPNG 82  
 QY 61 SLELPVAVGTDGEGIFRCQAMNNGKETKSNRYRVYQIPGKPEIVDSASELTAGVPKNVG 120  
 DB 83 SLELPVAVGTDGEGIFRCQAMNNGKETKSNRYRVYQIPGKPEIVDSASELTAGVPKNVG 142  
 QY 121 TCVSEGSYPAGTILSWHLDGKPLVPNEKGVSVKEQTRRHPEITGLTLOSELMTVPFARGSDP 180  
 DB 143 TCVSEGSYPAGTILSWHLDGKPLVPNEKGVSVKEQTRRHPEITGLTLOSELMTVPFARGSDP 202  
 QY 181 RPTESCSESGLPRKALRTAPLQPEVWEPVPLEVQLVPEEGAVAPGTVTLTCEVP 240  
 DB 203 RPTESCSESGLPRKALRTAPLQPEVWEPVPLEVQLVPEEGAVAPGTVTLTCEVP 262  
 QY 241 AQPSQTHMKTKGVPL 300  
 DB 263 AUPSFQIHWKMDGVPL 322  
 QY 301 PGREGPTAGSGGSGIG 317  
 DB 323 PGREGPTAGSGGSGIG 339

RESULT 11  
 AAE23218  
 ID AAE23218 standard: Protein. 416 AA  
 XX  
 AC AAE23218;  
 XX  
 DT 27-AUG-2002 (first entry)  
 XX  
 XX Cow receptor for advanced glycosylation end product (RAGE) protein.  
 XX  
 DE Cow: Receptor for advanced glycosylation end product; RAGE; cardiant;  
 XX tissue growth; neointimal formation; blood vessel, restenosis; diabetes;  
 KW myocardial infarction; angioplasty; peripheral vascular surgery; angina;  
 KW transgenic animal; acute thrombotic stroke, venous thrombosis.  
 XX  
 OS Ros taurus.  
 XX  
 FN WO200230889-A2.  
 XX  
 PD 18-APR-2002.  
 XX  
 XX 12-OCT-2001: 2001WO-US32336.  
 XX  
 PR 13-OCT-2000: 2000US-0687328.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Stern DM, Schmidt A, Marso S, Topol E, Lincoff AM;  
 XX  
 XX WPI: 2002-426260/45.  
 DR N-PSDB: AAD36951.  
 XX  
 PT Inhibiting new tissue growth or neointimal formation in blood vessels  
 PT of subject suffering from diabetes, stroke and preventing restenosis,  
 PT comprises administering inhibitor of receptor for advance glycation end  
 PT product -  
 XX  
 PS Disclosure: Page 14: 43pp: English.  
 XX  
 CC The invention relates to a method for inhibiting new tissue growth or  
 CC neointimal formation in blood vessels in a subject that has experienced  
 CC blood vessel injury and preventing exaggerated restenosis in a diabetic  
 CC subject. The method comprises administering an inhibitor of receptor for  
 CC advanced glycation/glycosylation end product (RAGE), so as to inhibit new  
 CC tissue growth or neointimal formation in subject's blood vessels and  
 CC preventing restenosis in the subject. The method is useful for inhibiting  
 CC new tissue growth or neointimal formation in blood vessels in a subject  
 CC like non-human animal, a transgenic non-human animal or a human suffering  
 CC from diabetes, acute thrombotic stroke, venous thrombosis, unstable  
 CC angina, myocardial infarction, abrupt closure following angioplasty or  
 CC stent placement, or thrombosis as a result of peripheral vascular surgery  
 CC The method is also useful for preventing restenosis and for determining  
 CC whether a compound inhibits new tissue growth in a blood vessel in a  
 CC subject. The present sequence is cow receptor for advanced glycosylation  
 CC end product (RAGE) protein.  
 XX  
 XX Sequence 416 AA;  
 XX  
 Query Match 82.6%; Score 1463; DB 23; Length 416;  
 Best Local Similarity 81.5%; Pred. No. 7.2e-102;  
 Matches 278; Conservative 21; Mismatches 30; Indels 12; Gaps 2;  
 QY 2 QNITARIGEPVLKCKGAPKPPORLEWKLNTGRTAEAKVLSPOGGPMDSVARVLPNGS 61  
 DB 24 QNITARIGEPVLKCKGAPKPPORLEWKLNTGRTAEAKVLSPOGGPMDSVARVLPNGS 82  
 QY 62 LELPVGITQDEGIFRCQAMNNGKETKSNRYRVYQIPGKPEIVDSASELTAGVPKNVG 121  
 DB 83 LLLPVGITQDEGIFRCQAMNNGKETKSNRYRVYQIPGKPEIVDSASELTAGVPKNVG 142  
 QY 122 CVSEGSYPAGTILSWHLDGKPLVPNEKGVSVKEQTRRHPEITGLTLOSELMTVPFARGSDP 181  
 DB 143 CVSEGSYPAGTILSWHLDGKPLVPNEKGVSVKEQTRRHPEITGLTLOSELMTVPFARGSDP 202



CC tissue growth or neointimal formation in subject's blood vessels and  
 CC preventing restenosis in the subject. The method is useful for inhibiting  
 CC new tissue growth or neointimal formation in blood vessels in a subject  
 CC like non-human animal, a transgenic non-human animal or a human suffering  
 CC from diabetes, acute thrombotic stroke, venous thrombosis, unstable  
 CC angina, myocardial infarction, abrupt closure following angioplasty or  
 CC stent placement, or thrombosis as a result of peripheral vascular surgery  
 CC The method is also useful for preventing restenosis and for determining  
 CC whether a compound inhibits new tissue growth in a blood vessel in a  
 CC subject. The present sequence is mouse receptor for advanced  
 CC glycosylation end product (RAGE) protein.

XX Sequence 403 AA;

Query Match 78.3%; Score 1483.5; DB 23; Length 403;  
 Best Local Similarity 79.6%; Pred. No. 6,5e-96;  
 Matches 262; Conservative 19; Mismatches 47; Indels 1; Gaps 1;  
 QY 2 QNITARIGEPVLVKCKGAPKPPQLEWKLNTGRTEAMKVLSPQGGPWDSVAPVLPNGS 61  
 Db 24 QNITARIGEPVLVLSCKGAPKPPQLEWKLNTGRTEAMKVLSPQ-GGPWDSVAPVLPNGS 82  
 QY 62 LFLFVAGSDERIFPQAMNPNKTKSNYPVRYQIPKPEIVDSASELTASVFNKVG 121  
 Db 83 LLLPATGIVDEGTFRCRATNRCKKSNYPVRYQIPKPEIVDPASELTASVFNKVG 142  
 QY 122 CVSEGSYPAGTILSWHLGKLLIPDGKELVKEETRRHPPTGLFTLPSLTVIPTGGTTH 181  
 Db 143 CVSEGSYPAGTILSWHLGKLLIPDGKELVKEETRRHPPTGLFTLPSLTVIPTGGTTH 202  
 QY 182 PTFSCSFPGLPHRALFTAPIQPRVMEPVLPVEEVLVVEDEGCAVACCCVITTCVPA 241  
 Db 203 PTFSCSFSGLPRLRRPLNTAPQLRVREPPEPGIQLVEPEGGIVAPGGVTLTCAISA 262  
 QY 242 QFSQIHWKMGVPELPEPSPVLLILPEIGSDGAGTYSVAIHSSHGPOESRAVSIIEP 301  
 Db 263 QPPQVHWIKDGAFLAPSPVLLILPEVGHADGEGTYSVATHPSHGCPSPPSVIRVET 322  
 QY 302 GREGPTAGSVGGSGTALALGILGLG 340  
 Db 323 GDEGPAEGSVGSGTALALGILGLG 351

RESULT 14

AAU77544

AAU77544 standard; Protein: 403 AA.

AAU77544;

05-JUN-2002 (first entry)

Murine receptor for advanced glycosylation end product (RAGE).

Receptor for advanced glycation end product; RAGE; receptor;  
 amyloid beta peptide; blood-brain barrier; neurovascular stress;  
 cerebral vasoconstriction suppressor; cerebral blood flow enhancer;  
 cerebral amyloid angiopathy; transgenic animal; amyloid beta precursor;  
 Alzheimer's disease; Down's syndrome; head trauma; stroke; mouse.

Mus musculus.

WO200214519-A1.

21-FEB-2002.

14-AUG-2001; 2001WO US25416.

14-AUG-2000; 2000US-0638648.

(UYCO ) UNIV COLUMBIA NEW YORK.

Stern DM, Schmidt AM, Yan SD, Zlokovic H;

DR

DR N-PSDB; ABK10857, ABK10858.

PT Ameliorating neurovascular stress and decreasing cerebral  
 PT vasoconstriction in subject suffering from chronic/acute cerebral  
 PT amyloid angiopathy, by administering inhibitor of receptor for advanced  
 PT glycation endproduct

PS Disclosure; Page 17-18; 68pp; English.

XX

CC The invention describes a method of ameliorating neurovascular stress,  
 CC and decreasing cerebral vasoconstriction in subject suffering from  
 CC chronic or acute cerebral amyloid angiopathy, comprising administering  
 CC an inhibitor (I) of receptor for advanced glycation end product (RAGE).  
 CC (I) inhibits transcytosis of amyloid beta peptides across blood-brain  
 CC barrier, thus decreasing cerebral vasoconstriction and increasing  
 CC cerebral blood flow. (I) is useful for treating amyloid angiopathy in a  
 CC subject, decreasing cerebral vasoconstriction in a transgenic non-human  
 CC animal (preferably, transgenic mouse overexpressing mutant human amyloid  
 CC beta precursor protein) or a human, suffering from chronic or acute  
 CC cerebral amyloid angiopathy, preferably, Alzheimer's disease, and for  
 CC ameliorating neurovascular stress comprising cerebral amyloid angiopathy  
 CC in a subject, where the neurovascular stress is caused by Alzheimer's  
 CC disease, aging, Down's syndrome, head trauma or stroke. This is the  
 CC amino acid sequence of murine receptor for advanced glycation end  
 CC product (RAGE) described in the invention.

XX Sequence 403 AA;

Query Match 78.3%; Score 1383.5; DB 23; Length 403;  
 Best Local Similarity 79.6%; Pred. No. 6,5e-96;  
 Matches 262; Conservative 19; Mismatches 47; Indels 1; Gaps 1;  
 QY 2 QNITARIGEPVLVKCKGAPKPPQLEWKLNTGRTEAMKVLSPQGGPWDSVAPVLPNGS 61  
 Db 24 QNITARIGEPVLVLSCKGAPKPPQLEWKLNTGRTEAMKVLSPQ-GGPWDSVAPVLPNGS 82  
 QY 62 LFLFVAGSDERIFPQAMNPNKTKSNYPVRYQIPKPEIVDSASELTASVFNKVG 121  
 Db 83 LLLPATGIVDEGTFRCRATNRCKKSNYPVRYQIPKPEIVDPASELTASVFNKVG 142  
 QY 122 CVSEGSYPAGTILSWHLGKLLIPDGKELVKEETRRHPPTGLFTLPSLTVIPTGGTTH 181  
 Db 143 CVSEGSYPAGTILSWHLGKLLIPDGKELVKEETRRHPPTGLFTLPSLTVIPTGGTTH 202  
 QY 182 PTFSCSFPGLPHRALFTAPIQPRVMEPVLPVEEVLVVEDEGCAVACCCVITTCVPA 241  
 Db 203 PTFSCSFSGLPRLRRPLNTAPQLRVREPPEPGIQLVEPEGGIVAPGGVTLTCAISA 262  
 QY 242 QFSQIHWKMGVPELPEPSPVLLILPEIGSDGAGTYSVAIHSSHGPOESRAVSIIEP 301  
 Db 263 QPPQVHWIKDGAFLAPSPVLLILPEVGHADGEGTYSVATHPSHGCPSPPSVIRVET 322

RESULT 15

AAU48747

ID AAM48747 standard; protein; 112 AA.

AAU48747;

02-APP-2002 (first entry)

Human RAGE protein V-domain SEQ ID NO 3.

Human: RAGE; receptor for advanced glycated endproduct; receptor;  
 anti-diabetic; neuroprotective; cytostatic; anti-inflammatory; vasolipic;  
 nephrotropic; dermatological; antiarteriosclerotic; nontropic; diabetes;  
 Alzheimer's disease; cancer; inflammation; kidney failure;  
 systemic lupus; nephritis; erectile dysfunction; atherosclerosis.

```

XX Homo sapiens
XX
XX WO200192892 A2.
XX
XX 06-DEC-2001.
XX
XX 40 MAY 2001; 2001WO US1/447.
XX
XX 30-MAY-2000; 2000US 207442P.
XX 05-MAR 2001; 2001US 079915Z.
XX
XX (TRAN ) TRANS TECH PHARMA.
XX
XX Shahbaz M.
XX
XX WPI; 2002 114372/15.
XX
XX Detecting a receptor for advanced glycated endproducts (RAGE)
XX modulators, for treating e.g., cancer, diabetes or inflammation,
XX comprises measuring the amount of bound anti-RAGE antibody
XX
XX Claim 1; Fig 2; 4pp; English.
XX
XX The invention relates to detecting receptor for advanced glycated
XX endproducts (RAGE) modulators comprises determining the amount of RAGE
XX protein or its fragment bound to the pre-adsorbed ligand by measuring the
XX amount of anti-RAGE antibody bound to the solid surface. The method is
XX useful for rapid, high throughput identification of compounds that
XX modulate RAGE. The compounds are useful for treating symptoms of diabetes
XX and symptoms of diabetic late complications, amyloidosis, Alzheimer's
XX disease, cancer, inflammation, kidney failure, systemic lupus nephritis
XX or inflammatory lupus nephritis, erectile dysfunction and
XX atherosclerosis.
XX
XX Sequence 112 AA:
XX
Query Match 44.7%; Score 595; Dk 24; Length 112;
Best Local Similarity 100.0%; Pred. No. 2.4e 37;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AQNITARIQEPLVLKTKGAHKPPQRLKWLNTGRTEAWKVLSPQGGPWDSVAVYLPNG 60
Db 1 AQNITARIQEPLVLKTKGAHKPPQRLKWLNTGRTEAWKVLSPQGGPWDSVAVYLPNG 60
QY 61 SLELPVAVGIQDEGIFRCQAMNNGKFKTSNYVPPVQIPSKPEIVDSASELT 112
Db 61 SLELPVAVGIQDEGIFRCQAMNNGKFKTSNYVPPVQIPSKPEIVDSASELT 112

```

Search completed: May 30, 2003, 15:55:05  
Job time : 76.5225 secs



Copyright (c) 1993 - 2003 CompuGen Inc

OM protein - protein search, using sw model

Run on: May 30, 2003, 15:55:12, Search time 26.9189 seconds  
(without alignments)  
362.883 Million cell updates/sec

Title: US-09-872 185b 2  
Perfect score: 1766  
Sequence: 1 AQNITAR:GPIVLKCKGAPKPPORLEWKLNTGRTEAMKVLSPGGGHWDSVARVLPNG 332

Scoring table: BLOSUM62

Gap: 10 0 0 Gap: 10 0 0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: 1: /cgn2\_6/ptedata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptedata/1/iaa/5A\_COMB.pep.\*  
3: /cgn2\_6/ptedata/1/iaa/5A\_COMB.pep.\*  
4: /cgn2\_6/ptedata/1/iaa/5A\_COMB.pep.\*  
5: /cgn2\_6/ptedata/1/iaa/5A\_COMB.pep.\*  
6: /cgn2\_6/ptedata/1/iaa/5A\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1766	100.0	332	US-09-062-365-1	Sequence 1, Appli
2	1701	96.3	318	US-08-633-148-4	Sequence 4, Appli
3	1701	96.3	340	US-08-633-148-2	Sequence 2, Appli
4	1495	84.7	278	US-08-432-016-5	Sequence 5, Appli
5	1495	84.7	278	US-08-684-594-5	Sequence 5, Appli
6	218.5	12.4	583	US-08-432-016-2	Sequence 2, Appli
7	218.5	12.4	583	US-08-684-594-2	Sequence 2, Appli
8	212.5	12.0	456	US-08-432-016-4	Sequence 4, Appli
9	212.5	12.0	456	US-08-684-594-4	Sequence 4, Appli
10	209	11.8	1653	US-09-540-245A-18	Sequence 18, Appli
11	205	11.6	477	US-08-432-016-3	Sequence 3, Appli
12	205	11.6	477	US-08-684-594-3	Sequence 3, Appli
13	197	11.2	486	US-08-432-016-6	Sequence 6, Appli
14	197	11.2	486	US-08-684-594-6	Sequence 6, Appli
15	190.5	10.8	698	US-08-602-725-16	Sequence 36, Appli
16	190.5	10.8	734	US-08-389-459A-17	Sequence 17, Appli
17	190.5	10.8	744	US-08-987-867A-17	Sequence 17, Appli
18	189.5	10.7	1297	US-09-540-245A-17	Sequence 17, Appli
19	177	10.0	924	US-08-481-150-28	Sequence 28, Appli
20	177	10.0	924	US-08-656-984A-28	Sequence 28, Appli
21	177	10.0	924	US-08-481-150-28	Sequence 28, Appli
22	177	10.0	924	US-08-481-150-28	Sequence 28, Appli
23	176.5	10.0	421	US-08-659-984A-1	Sequence 1, Appli
24	176.5	10.0	421	US-08-660-531-1	Sequence 1, Appli
25	176.5	10.0	444	US-08-659-984A-5	Sequence 5, Appli
26	176.5	10.0	444	US-08-660-531-5	Sequence 5, Appli
27	172.5	9.8	642	US-08-217-299-1	Sequence 1, Appli

28	172	9.7	1266	4	US-08-506-296B-4	Sequence 4, Appli
29	170	9.6	462	2	US-08-752-307B-7	Sequence 7, Appli
30	170	9.6	462	4	US-09-707-802-7	Sequence 7, Appli
31	170	9.6	462	4	US-09-941-426-7	Sequence 7, Appli
32	170	9.6	465	2	US-08-752-307B-5	Sequence 5, Appli
33	170	9.6	465	4	US-09-707-802-5	Sequence 5, Appli
34	170	9.6	465	4	US-09-941-426-5	Sequence 5, Appli
35	168	9.5	1095	4	US-09-510-245A-15	Sequence 15, Appli
36	165	9.3	607	2	US-08-752-307B-12	Sequence 12, Appli
37	165	9.3	607	4	US-09-707-802-12	Sequence 12, Appli
38	165	9.3	607	4	US-09-941-426-12	Sequence 12, Appli
39	164.5	9.4	1094	4	US-08-481-150-28	Sequence 28, Appli
40	164	9.3	1447	4	US-09-041-886-25	Sequence 25, Appli
41	164	9.3	1447	5	US-09-041-886-25	Sequence 25, Appli
42	162	9.2	30	4	US-09-062-365-5	Sequence 5, Appli
43	159.5	9.0	1101	3	US-08-586-485-2	Sequence 2, Appli
44	158.5	9.0	1253	4	US-08-506-296B-14	Sequence 14, Appli
45	158	8.9	612	2	US-08-752-307B-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-09-062-365-1  
: Sequence 1, Application US/J9062365  
: Patent No. 6465422  
: GENERAL INFORMATION:  
: APPLICANT: Schmidt, Ann Marie  
: APPLICANT: Stern, David  
: TITLE OF INVENTION: METHOD FOR INHIBITING TUMOR INVASION OR SPREADING IN A  
: TITLE OF INVENTION: SUBJECT  
: REFERENCE: 55124  
: CURRENT APPLICATION NUMBER: US/09/062,365  
: CURRENT FILING DATE: 1998-04-17  
: NUMBER OF SEQ ID NOS: 6  
: SOFTWARE: PatentIn Ver. 2.1  
: SEQ ID NO 1  
: LENGTH: 332  
: TYPE: PRT  
: ORGANISM: Human  
US-09-062-365-1

Query Match	100.0%	Score	1766	DB	4	Length	332
Best Local Similarity	100.0%	Pred. No.	1.2e-141				
Matches	332	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Q1	1	AQNITAR	GPIVLKCKGAPKPPORLEWKLNTGRTEAMKVLSPGGGHWDSVARVLPNG	60			
D1	1	AQNITAR	GPIVLKCKGAPKPPORLEWKLNTGRTEAMKVLSPGGGHWDSVARVLPNG	60			
C1	61	SLFLPVGIGDEIFRCQAMNRNGKETKSNKYRVYVQIPGKPEIVDSASELTAGVKNVG	120				
D1	61	SLFLPVGIGDEIFRCQAMNRNGKETKSNKYRVYVQIPGKPEIVDSASELTAGVKNVG	120				
Q1	121	TCVSEGSYPAGTISWELDKCKPLVPNEKYSVVKRPHPHETGIFTLQSHLWTFAGGHP	180				
D1	121	TCVSEGSYPAGTISWELDKCKPLVPNEKYSVVKRPHPHETGIFTLQSHLWTFAGGHP	180				
Q1	181	RPTFSFSPSGIPLPHALPTAPIQPPVWEPVPLFVQIVVEPFGCAVAGCTVTITCEVP	240				
D1	181	RPTFSFSPSGIPLPHALPTAPIQPPVWEPVPLFVQIVVEPFGCAVAGCTVTITCEVP	240				
Q1	241	AQTSFQIIMWPIVVEPFLFSPVLLIPELFGVQNTYSVATISSHHPESPASVISTIE	300				
D1	241	AQTSFQIIMWPIVVEPFLFSPVLLIPELFGVQNTYSVATISSHHPESPASVISTIE	300				
Q1	301	FGREGTADSSVSSGIGLALAGTICGLTIA	332				
D1	301	FGREGTADSSVSSGIGLALAGTICGLTIA	332				

RESULT 2

```

US 08 644 148 4
: Sequence 4, Application US/08644148
: Patent No. 5864018
: GENERAL INFORMATION:
: APPLICANT: MORSE, MICHAEL J.
: APPLICANT: NAGASHIMA, MARIKO
: APPLICANT: HOLLANDER, DORIS A.
: TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
: TITLE OF INVENTION: END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TOWNSEND & TOWNSEND & CROW LLP
: STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
: CITY: SAN FRANCISCO
: STATE: CALIFORNIA
: COUNTRY: U.S.A.
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC DOS/MS DOS
: SOFTWARE: Patent In Release #1.0, Version #1.40
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08644148
: FILING DATE: 16 APR 1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: MURPHY ESQ., MATTHEW B.
: REGISTRATION NUMBER: 39,787
: REFERENCE/DOCKET NUMBER: 014618-00560005
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 326-2400
: TELEFAX: (415) 326-2422
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 418 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US 08-644-148 4

```

```

Query Match 96.58; Score 1701; DB 2; Length 418;
Best local Similarity 100.0%; Pred. No. 3,70-146;
Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTTARIGEPVLVKKAPKKEPQLEWKLNTGKTEAWKVLSPGSGGWDSVARVLPNG 60
DB 1 AONTTARIGEPVLVKKAPKKEPQLEWKLNTGKTEAWKVLSPGSGGWDSVARVLPNG 60

QY 61 SLFLPAGVLIQDEGTFEFCQAMNNPKETKSNYRVVYQIPGKPEIVDSASELTASVFNKVG 120
DB 61 SLFLPAGVLIQDEGTFEFCQAMNNPKETKSNYRVVYQIPGKPEIVDSASELTASVFNKVG 120

QY 121 TCVSEGSYPACILSWHLKPKPLVNEKGVSKQETRRHPDTGLFTLQSELMTVPARGGDP 180
DB 121 TCVSEGSYPACILSWHLKPKPLVNEKGVSKQETRRHPDTGLFTLQSELMTVPARGGDP 180

QY 181 RPTFSCSFSPGLPRHRLFTAPIQPRVWEPVPLEEVQIVVEPGGAVAPGVTITCEVP 240
DB 181 RPTFSCSFSPGLPRHRLFTAPIQPRVWEPVPLEEVQIVVEPGGAVAPGVTITCEVP 240

QY 241 AQPSPQIHWKKGVPLPLPSPVLLPEIGPQDQKTSVATIHSSHGQESRAVSTSTIE 400
DB 241 AQPSPQIHWKKGVPLPLPSPVLLPEIGPQDQKTSVATIHSSHGQESRAVSTSTIE 400

QY 401 PGEGGPIAGSVGGSGIGT 418
DB 401 PGEGGPIAGSVGGSGIGT 418

```

```

RESULT 4
US-08-644-148 2
: Sequence 4, Application US/08644148

```

```

: Sequence 2, Application US/08644148
: Patent No. 5864018
: GENERAL INFORMATION:
: APPLICANT: MORSE, MICHAEL J.
: APPLICANT: NAGASHIMA, MARIKO
: APPLICANT: HOLLANDER, DORIS A.
: TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION
: TITLE OF INVENTION: END-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: TOWNSEND & TOWNSEND & CROW LLP
: STREET: TWO EMBARCADERO CENTER, 8TH FLOOR
: CITY: SAN FRANCISCO
: STATE: CALIFORNIA
: COUNTRY: U.S.A.
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC DOS/MS DOS
: SOFTWARE: Patent In Release #1.0, Version #1.40
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08644148
: FILING DATE: 16 APR 1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: MURPHY ESQ., MATTHEW B.
: REGISTRATION NUMBER: 39,787
: REFERENCE/DOCKET NUMBER: 014618-00560005
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 326-2400
: TELEFAX: (415) 326-2422
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 340 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US 08-644-148-2

```

```

Query Match 96.58; Score 1701; DB 2; Length 340;
Best local Similarity 100.0%; Pred. No. 40-146;
Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTTARIGEPVLVKKAPKKEPQLEWKLNTGKTEAWKVLSPGSGGWDSVARVLPNG 60
DB 1 AONTTARIGEPVLVKKAPKKEPQLEWKLNTGKTEAWKVLSPGSGGWDSVARVLPNG 60

QY 61 SLFLPAGVLIQDEGTFEFCQAMNNPKETKSNYRVVYQIPGKPEIVDSASELTASVFNKVG 120
DB 61 SLFLPAGVLIQDEGTFEFCQAMNNPKETKSNYRVVYQIPGKPEIVDSASELTASVFNKVG 120

QY 121 TCVSEGSYPACILSWHLKPKPLVNEKGVSKQETRRHPDTGLFTLQSELMTVPARGGDP 180
DB 121 TCVSEGSYPACILSWHLKPKPLVNEKGVSKQETRRHPDTGLFTLQSELMTVPARGGDP 202

QY 181 RPTFSCSFSPGLPRHRLFTAPIQPRVWEPVPLEEVQIVVEPGGAVAPGVTITCEVP 240
DB 181 RPTFSCSFSPGLPRHRLFTAPIQPRVWEPVPLEEVQIVVEPGGAVAPGVTITCEVP 262

QY 241 AQPSPQIHWKKGVPLPLPSPVLLPEIGPQDQKTSVATIHSSHGQESRAVSTSTIE 400
DB 263 AQPSPQIHWKKGVPLPLPSPVLLPEIGPQDQKTSVATIHSSHGQESRAVSTSTIE 422

QY 401 PGEGGPIAGSVGGSGIGT 418
DB 423 PGEGGPIAGSVGGSGIGT 440

```

```

RESULT 4
US 08 442-016-5
: Sequence 5, Application US/08442016

```

Patent No. 5968768  
 GENERAL INFORMATION:  
 APPLICANT: HAYNES, BARTON F.  
 APPLICANT: APOFFO, ALEANDRO  
 APPLICANT: PATEL, DHAVALKUMAR  
 APPLICANT: BOWEN, MICHAEL A.  
 APPLICANT: MARQUARDT, HANS  
 TITLE OF INVENTION: CD6 LIGAND  
 NUMBER OF SEQUENCES: 9  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: NIXON & VANDERHYE P.C.  
 STREET: 1100 NORTH GLEBE ROAD  
 CITY: ARLINGTON  
 STATE: VIRGINIA  
 COUNTRY: U.S.A.  
 ZIP: 22201-4714  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/432,016  
 FILING DATE: 01-MAY-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/333,350  
 FILING DATE: 02-NOV-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/143,903  
 FILING DATE: 02-NOV-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: WILSON, MARY J.  
 REGISTRATION NUMBER: 32,955  
 REFERENCE/DOCKET NUMBER: 1579-95  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 816-4000  
 TELEFAX: (703) 816-4100  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 278 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-432-016-5

Query Match 84.7%; Score 1495; DB 2; Length 278;  
 Best Local Similarity 99.6%; Pred. No. 8e-119;  
 Matches 277; Conservative 1; Mismatches 0; Indels 0; Gaps 0.

QY 8 IGEPLVLKCKGAKKKPKPPQRLKWKLNTRTEAMKVLSPGGGPDWSVAVPLPNCGLFLPAV 67  
 DB 1 IGEPLVLKCKGAKKKPKPPQRLKWKLNTRTEAMKVLSPGGGPDWSVAVPLPNCGLFLPAV 60

QY 68 GIQDGGIFPRQAMNPNKSTKSNYPVRYQIPCKPEIVDSASFLTAGVPNKVSTCVSEGS 127  
 DB 61 GIQDGGIFPRQAMNPNKSTKSNYPVRYQIPCKPEIVDSASFLTAGVPNKVSTCVSEGS 120

QY 128 YPAGTLSWHLDDGKPLVFNKGVSKVKEQTRRHPTGLFTIQSELMTVPARGSDPRPTFSCS 187  
 DB 121 YPAGTLSWHLDDGKPLVFNKGVSKVKEQTRRHPTGLFTIQSELMTVPARGSDPRPTFSCS 180

QY 188 FSPGLPRHRLRATAPQPRVWEPVPLEEVQLVVEPESGAVAPGGIVILICRVPAPQPSQI 247  
 DB 181 FSPGLPRHRLRATAPQPRVWEPVPLEEVQLVVEPESGAVAPGGIVILICRVPAPQPSQI 240

QY 248 HMKDGVPLPLPSPVLLPILPGLPQGGTYSCVATHSS 285  
 DB 241 HMKDGVPLPLPSPVLLPILPGLPQGGTYSCVATHSS 278

RESULT 5

US-08-684-594-5  
 Sequence 5, Application US/08684594  
 Patent No. 5998172  
 GENERAL INFORMATION:  
 APPLICANT: HAYNES, BARTON F.  
 APPLICANT: APOFFO, ALEANDRO  
 APPLICANT: PATEL, DHAVALKUMAR  
 APPLICANT: BOWEN, MICHAEL A.  
 APPLICANT: MARQUARDT, HANS  
 TITLE OF INVENTION: CD6 LIGAND  
 NUMBER OF SEQUENCES: 14  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: NIXON & VANDERHYE P.C.  
 STREET: 1100 NORTH GLEBE ROAD  
 CITY: ARLINGTON  
 STATE: VIRGINIA  
 COUNTRY: U.S.A.  
 ZIP: 22201-4714  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/684,594  
 FILING DATE: 18-JUL-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/432,016  
 FILING DATE: 01-MAY-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/333,350  
 FILING DATE: 02-NOV-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/143,903  
 FILING DATE: 02-NOV-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: WILSON, MARY J.  
 REGISTRATION NUMBER: 32,955  
 REFERENCE/DOCKET NUMBER: 1579-112  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 816-4000  
 TELEFAX: (703) 816-4100  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 278 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-684-594-5

Query Match 84.7%; Score 1495; DB 2; Length 278;  
 Best Local Similarity 99.6%; Pred. No. 8e-119;  
 Matches 277; Conservative 1; Mismatches 0; Indels 0; Gaps 0.

QY 8 IGEPLVLKCKGAKKKPKPPQRLKWKLNTRTEAMKVLSPGGGPDWSVAVPLPNCGLFLPAV 67  
 DB 1 IGEPLVLKCKGAKKKPKPPQRLKWKLNTRTEAMKVLSPGGGPDWSVAVPLPNCGLFLPAV 60

QY 68 GIQDGGIFPRQAMNPNKSTKSNYPVRYQIPCKPEIVDSASFLTAGVPNKVSTCVSEGS 127  
 DB 61 GIQDGGIFPRQAMNPNKSTKSNYPVRYQIPCKPEIVDSASFLTAGVPNKVSTCVSEGS 120

QY 128 YPAGTLSWHLDDGKPLVFNKGVSKVKEQTRRHPTGLFTIQSELMTVPARGSDPRPTFSCS 187  
 DB 121 YPAGTLSWHLDDGKPLVFNKGVSKVKEQTRRHPTGLFTIQSELMTVPARGSDPRPTFSCS 180

QY 188 FSPGLPRHRLRATAPQPRVWEPVPLEEVQLVVEPESGAVAPGGIVILICRVPAPQPSQI 247  
 DB 181 FSPGLPRHRLRATAPQPRVWEPVPLEEVQLVVEPESGAVAPGGIVILICRVPAPQPSQI 240

QY 248 HMKDGVPLPLPSPVLLPILPGLPQGGTYSCVATHSS 285

us-09-872-185b-2.ra1

Fri May 30 17:16:36 2003

```

160 ETGLFTLOSRLMVTYPARGGHPPIESST-ESQGLPHRAIRGALQHRWME IVDLRE 215
195 VTQLYTMISTLEYKTKR-ADIQMPFETSVLYGFS-VGKTHSEGVAVDIAVYPTEQ 248
216 VQLVVEPEGAVAPGGHVTITTEVEFAQSP-QIHMNKDGVFLPDPSPVLLILPEIGPQIQ 274
249 VTIOVLPPKKAKEGUNIT-KGLNGNPPHEDFLEFLPQGPQHCIRSSNYITLXIVRRNAT 408
275 GTYSC 279
309 GDYKC 313

RESULT 7
US-08-684-594-2
; Sequence 2, Application US/08684594
; Patent No. 5958172
; GENERAL INFORMATION:
; APPLICANT: HAYNES, BARTON F.
; APPLICANT: ARUFFO, ALEJANDRO
; APPLICANT: PATEL, DHAVALKUMAR
; APPLICANT: BOWEN, MICHAEL A.
; APPLICANT: MARQUARDT, HANS
; TITLE OF INVENTION: CD6 LIGAND
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432-016
; FILING DATE: 01-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/433,450
; FILING DATE: 02-NOV-1994
; APPLICATION NUMBER: US 08/433,450
; FILING DATE: 02-NOV-1994
; APPLICATION NUMBER: US 08/143,903
; FILING DATE: 02-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579 112
; TELEPHONE: (703) 816,4000
; TELEFAX: (703) 816,4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; OTHER INFORMATION: Xaa - M/S
; OTHER INFORMATION: Xaa - M/T
; US-08-684-594-2

Query Match 12.4%; Score 218.5; DB 2; Length 583;
Best Local Similarity 25.9%; Pred. No. 1.9e-10;
Matches 79; Conservative 45; Mismatches 130; Indels 51; Gaps 12;

4 ITARIGRPIVICKGAPKPKPQRLKWLNTGRTKAKVLSPOGGG-----PW 50
41 VNSAYGDTLIPGR--LDVQNLMP---GK---WKYEKPDGSEVFIAFRSSTKKSQY 80
51 DSV-----ARVLNPGSLFLPAVGIDQGLIFRCQANRNCKVTKSNY-----RVAVYQIP 99
81 DDDVPEYKDRINLSINYFLISNARIISDERKRFVCMV-----TEDNVFAPTIKVFQKP 134
100 GKPEIVDSASELTAGVFNKVGTCVSEGSYPACTLSWILDGKPLVFNKGVSVKQETRRHP 159
135 SKPEIVSKAFLTELEQLKQKADGCSISELSYDNGNTWYRNCKVTHDLEGAVALIPKKEMDP 194

```

31 VNSAYGDTIIIPCP---LVDQNLMF---GK---WKYEKPSPEVIAFPSTKKSVMY 80  
Db  
51 DSV ARVLIMGSLFPAVGQIQDGIIRGQAMNRNKKETPKSNY----RVRVYQIP 99  
QY  
81 DDVPEYKQRLNLSNYTLSTISNARISDEKFEVCLV-----TEDNVFEAPTIYKVKFQ 134  
Db  
100 AKPFTVTSASSLIAVFNKVGIVWSESYSGATLSWHLEKPLVPVNEKIVSYKREATPPRP 159  
QY  
135 SKPEIVSKALELEINQKKKQDGLSELSYSPGNIIMYRNKGVHPLLEGAAVILFKKEMLP 194  
Db  
160 EPTGLTQSLQVMTPARGQGRPPPTFSY---FSPGIPPHPALDPAIQPVME-PVPLEE 215  
QY  
195 VTQLTMTSTLEYKITK-ADIQMPETCSVTYAPS-----GQRTIISEQNVDEIYYPTSQ 248  
Db  
276 VOLVVERPSCAAVAGGVVILPQWTAQIST-QIHWKKGQVILPILFSPVILFELGPDQC 274  
QY  
249 VTIVLPPKXAIKEGNIITLCKLGNRNPPPEEFELYPQAPPEIFISNTYITLXGVFNAL 308  
Db  
275 RYSG 279  
QY  
309 GDYK 313  
Db

## RESULT 8

US-08-432-016-4

```

: Sequence 4. Application US/04432016
: Patent No. 5968768
: GENERAL INFORMATION
: APPLICANT: HAYNES, BARTON F.
: APPLICANT: ARUFFO, ALEJANDRO
: APPLICANT: PATEL, CHAVALKUMAR
: APPLICANT: BOWEN, MICHAEL A.
: APPLICANT: MARQUARDT, HANS
: TITLE OF INVENTION: CUB LIGAND
: NUMBER OF SEQUENCES: 9
: CORRESPONDENT'S ADDRESS
: ADDRESSEE: NIXON & VANDERHYE P.C.
: STREET: 1100 NORTH GLEBE ROAD
: CITY: ARLINGTON
: STATE: VIRGINIA
: COUNTRY: U.S.A.
: ZIP: 22201-4714
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1 0,
: CURRENT APPLICATION DATA
: APPLICATION NUMBER: US/08/432,016
: FILING DATE: 01-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/333,450
: FILING DATE: 02 NOV 1994
: PRIOR APPLICATION NUMBER:
: APPLICATION NUMBER: US 08/143,903
: FILING DATE: 02-NOV-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: WILSON, MARY J.
: REGISTRATION NUMBER: 32,955
: REFERENCE/DOCKET NUMBER: 1579-95
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 816-4000
: TELEFAX: (703) 816-4100
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 466 amino acids
: TYPE: amino acid
: STRANDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-432-016-4

```

## RESULTS

US-08/684-594-4  
Sequence 4, Application US/08684594  
Patent No. 5998172  
GENERAL INFORMATION:  
APPLICANT: HAYNES, BARTON F.  
APPLICANT: ARUFFO, ALEJANDRO  
APPLICANT: PATEL, DHAVALKUMAR  
APPLICANT: BOWEN, MICHAEL A.  
APPLICANT: MARQUARDT, HANS  
TITLE OF INVENTION: CD6 LIGAND  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERSHYE P.C.  
STREET: 1100 NORTH GLADE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS DOS  
SOFTWARE: Patent In Release #1.0, V  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/684,594  
FILING DATE: 18-JUL-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/432,016  
FILING DATE: 01-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/333,350  
FILING DATE: 02-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/143,903  
FILING DATE: 02-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: WILSON, MARY J.  
REGISTRATION NUMBER: 72,955  
REFERENCE/DOCKET NUMBER: 1579-112  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 466 amino acids

```

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
DS 08-684-594-4

Query Match      12.0%; Score 212.5; DB 2; Length 466;
Best local similarity 25.5%; Pred. No. 4.6e-10;
Matches 76; Conservative 40; Mismatches 133; Indels 49; Gaps 11;

          9  GEPVLVKKCAVPKKPKPQLEKWLKNTGPTCAKKVLSPOQGGHWSVAR ----- 55
          11  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          2  GETIVVPCNGETKKPKGLIFK -----KKYVK DQSSPGDLLVQAQKDEATVSATD 52
          11  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          56  VLFNGSLFLPAVGHQDEGIFPCQAMNENKFKTSNYPVRYVQLPKKDEIVDSA 108
          11  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          53  GYKSRVSI AANSSLLAQGSLADQKVFCTMVVSPFNLEYS-VLVKVKHKPSAIVIKNA 111
          11  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          109  SETAGVPNKKVGTGVSSEGYSPACTLSWILDKKPLVPNEKGVSVKQFTRIPETGLTLOS 168
          11  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          112  KELENGKLTQIQGKGVVENANPADLLWKNNQTVLDGKTLTSTTKKIKIGLSSTSS 171
          11  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          169  ELAVTVARGQDDEPTFSFSFGLIKRKAIRKTAIPQFVWEP -----VVLVQGVWEP 222
          11  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          172  ELQVT-ARKEDVESQTET- AKH VNGPQC- VSEFESITTHYTEKKSQLOVVS 250
          11  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          223  EGGAVALGGTHTVITGCPVAPSPD-QTHMKKDGVPILPSPSVLLPELGGDQGTYS 279
          11  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          221  Q-SIPRPGHIVTLKCAKGNPPHPSNFNFKGKKVTVTKIVYTLTGVTADSGIYK 277
          11  | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 10
US-09-540-245A-18
: Sequence 18, Application US/09/040245A
: Patent No. 6270984
:
: GENERAL INFORMATION:
: APPLICANT: Goodman, Corey
: APPLICANT: Kid, Thomas
: APPLICANT: Brose, Katja
: APPLICANT: Tessier Lavigne, Marc
: TITLE OF INVENTION: Modularizing Robo: Lizard Interactions
: FILE REFERENCE: B98-031 4
: CURRENT APPLICATION NUMBER: US/09/540,245A
: CURRENT FILING DATE: 2000-03 31
: PRIOR APPLICATION NUMBER: 60/065,544
: PRIOR FILING DATE: 1997 11 14
: PRIOR APPLICATION NUMBER: 60/081,057
: PRIOR FILING DATE: 1998 04 07
: NUMBER OF SEQ ID NOS: 20
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 18
: LENGTH: 1651
: TYPE: PRT
: ORGANISM: human
US-09-540-245A-18

```

[illegible]

QY 179 DRRITSCSPSLPHEALPTAPICQVWE-PVPLEEVQVLPVPECCAVAGCIVTLIC 237  
 Db 179 DANAKFTCIYVYHGPSQ--KTUSEPVFVIVPTEKVTIRVLSQSSTIKEDGNVTLK 236  
 QY 238 EVPAQSPQ----- 246  
 Db 237 SCNCNPPPPQFFLYIPGTEBIRSSDTYVYVMTDVPNNATGEYKCSLLDKSMDDTTITVHY 296  
 QY 247 -----IHWKKGVPPLPLPSPVLLPEIGPQDO 274  
 Db 237 LDGLQITSCGEVTKIGALEVVSATISSSNAIVFWIKNTKMTSIS--FSSLOYQDA 352  
 QY 275 GTYSVATHSHGHQESKAVSISIE 300  
 Db 353 GNYICETTHKEVEGLKKRKLKIVE 378

RESULT 12  
 US-08-684-594-3  
 ; Sequence 1, Application US/08684534  
 ; Patent No. 5998172  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HAYNES, BARTON F.  
 ; APPLICANT: ARUFFO, ALEJANDRO  
 ; APPLICANT: PATEL, DHAVALKUMAR  
 ; APPLICANT: BOWEN, MICHAEL A.  
 ; APPLICANT: MARQUARDT, HANS  
 ; TITLE OF INVENTION: CD6 LIGAND  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: NIXON & VANDEPHYE, P.C.  
 ; STREET: 1100 NORTH GL-HE ROAD  
 ; CITY: ARLINGTON  
 ; STATE: VIRGINIA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 22201-4714  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/684,594  
 ; FILING DATE: 18-JUL-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/432,016  
 ; FILING DATE: 01-MAY-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/333,350  
 ; FILING DATE: 02-NOV-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/143,903  
 ; FILING DATE: 02-NOV-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: WILSON, MARY J.  
 ; REGISTRATION NUMBER: 32,955  
 ; REFERENCE/DOCKET NUMBER: 1579-112  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 816-4000  
 ; TELEFAX: (703) 816-4100  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 477 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-684-594-3

Query Match 11.6%; Score 205; DB 2; Length 477.  
 Best Local Similarity 21.5%; Proc. No. 2e-09;

Matches 70; Conservative 39; Mismatches 125; Indels 92; Gaps 6.  
 QY 59 NGSLLFAVIGLQDE-IFRQJAMNKN:KETKSNYPVAVYLIPOKPEIVDSASSETAGVPMK 118  
 Db 61 NYLLISINARISDEKREVCMLVTEDDVSEPTV-KVFKQPSQPEILHQADELETEKLM 119  
 QY 119 VGTCSGSGSPAGTSLSHLDGKPLVPNEKCVSKQTRRHPTETGLFTLOSRLMVTIPARGG 178  
 Db 120 LGEVWDSYPEGNVNTWKNGRVLPQVEEVVNLNRKVENRSTGLFTMTSSLSQYMPK-E 178  
 QY 179 DPPFTSCSPSLPPIPAIPATIPAPVWE-PVPLEEVQVLPVPECCAVAGCIVTLIC 237  
 Db 179 DANAKFTCIYVYHGPSQ--KTUSEPVFVIVPTEKVTIRVLSQSSTIKEDGNVTLK 236  
 QY 238 EVPAQSPQ----- 246  
 Db 237 SCNCNPPPPQFFLYIPGTEBIRSSDTYVYVMTDVPNNATGEYKCSLLDKSMDDTTITVHY 296  
 QY 247 -----IHWKKGVPPLPLPSPVLLPEIGPQDO 274  
 Db 237 LDGLQITSCGEVTKIGALEVVSATISSSNAIVFWIKNTKMTSIS--FSSLOYQDA 352  
 QY 275 GTYSVATHSHGHQESKAVSISIE 300  
 Db 353 GNYICETTHKEVEGLKKRKLKIVE 378

RESULT 13  
 US-08-432-016-6  
 ; Sequence 6, Application US/08432016  
 ; Patent No. 598768  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HAYNES, BARTON F.  
 ; APPLICANT: ARUFFO, ALEJANDRO  
 ; APPLICANT: PATEL, DHAVALKUMAR  
 ; APPLICANT: BOWEN, MICHAEL A.  
 ; APPLICANT: MARQUARDT, HANS  
 ; TITLE OF INVENTION: CD6 LIGAND  
 ; NUMBER OF SEQUENCES: 9  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: NIXON & VANDEPHYE, P.C.  
 ; STREET: 1100 NORTH GL-HE ROAD  
 ; CITY: ARLINGTON  
 ; STATE: VIRGINIA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 22201-4714  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/432,016  
 ; FILING DATE: 01-MAY-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/333,350  
 ; FILING DATE: 02-NOV-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/143,903  
 ; FILING DATE: 02-NOV-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: WILSON, MARY J.  
 ; REGISTRATION NUMBER: 32,955  
 ; REFERENCE/DOCKET NUMBER: 1579-95  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 816-4000  
 ; TELEFAX: (703) 816-4100  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 486 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:

[illegible]

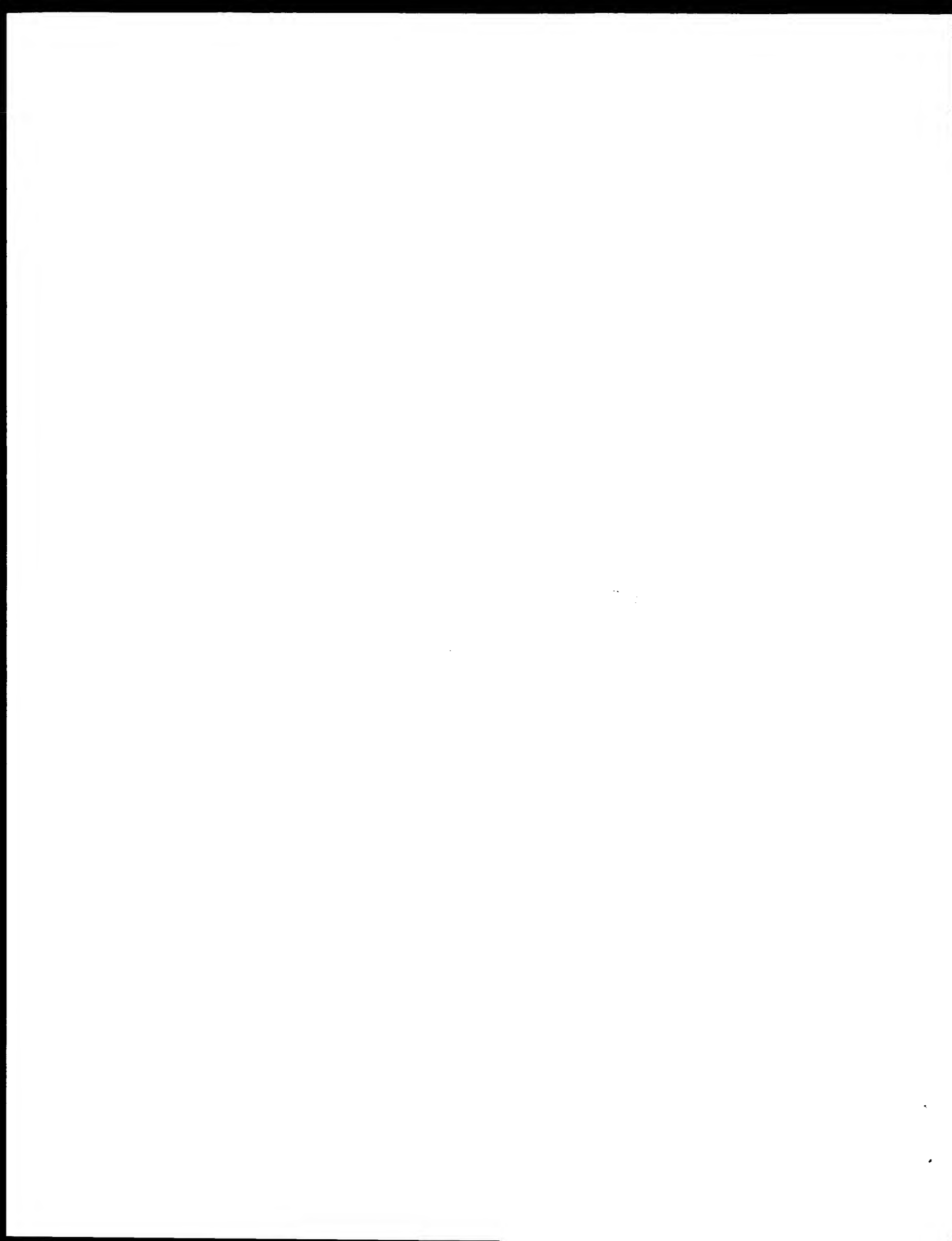


```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/602,725
FILING DATE: 02-FEB-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB94/01816
FILING DATE: 19-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9317423
FILING DATE: 21-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36663
REFERENCE/DECKET NUMBER: 1060-8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4091
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 698 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-602-725-36

Query Match      10.88; Score 190.5; DB 2; Length 698;
Best local Similarity 17.44; Prod No 5.8e-08;
Matches 89; Conservative 62; Mismatches 134; Indels 21; Gaps 16.

QY      3 NITAPIGEPLVKCKAPKPPQPPQLEWKLNTGTEAWKVLSPQGGGPPQDSVAVPLPNSGL 62
DB      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      63 FLAVAGLQDRGIEKQAMNEN-CKEIKSNYRVKVYQIPCKPIVDSASE----- 110
DB      1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      281 FIPNITVNSGSYTCQAHNSDTGLNPTTITVYAFPPKPTITSNNSNPVEDDAVALI 340
QY      111 -----LI-----AGVFNKVGTCVS 124
DB      341 CEPQLQNTVYIWWNVNGLIVSPHPLQLSDNINRTLLLSVTRNVGVYEGGIONELSDVHS 400
QY      125 E-----GSPAGTLSMHLGDKPLVPNEKGV 150
DB      401 DPVILNVLYGDDPTISPSYTYRPGVNLSSLCHAASNPAPQYSLIDG-----N 450
QY      151 VKQTRR-----HPTGIEFTLQSFIMVTPAGGH-----PRTFSCSPSP 190
DB      451 IQHTQELFISNTERNSGLYTCQAN---NSASGHSRTTVKTTVTSaelPRPSISSNSK 507
QY      191 GLPRHAL-----FTAPIQPPVW----- 208
DB      508 PVHUKDAVATQEPKAGNFIYIWWNVGSHIPYSPKPIQLSNKPKLITFNVTINAPAYVC 557
QY      209 -----EPVPLEEV-----QIVVFPFGGAVADG3TVTICEVPAGPSPQIHWMDG 253
DB      568 GIQNSVANSDDVTLVLYGPTPIISPPDSYLSGANLNSCHISANSPSQYSWRING 627
QY      254 VPLPPLSPVLIILPEIGPQGGIYSCVAIHSSHGPGCFSPKAVSISIEPGEFGPTAGSVGG 313
DB      628 I--PQHTQVLFTAKITPNNNGTYACFVSNATGRNNSTIVKSTIVSASGT-----SPGL 679
QY      314 SGLGTALALGILGGL 329
DB      680 SACATVGTIMIGVILGV 695
```

Search completed: May 30, 2003, 15:59:51  
Job time : 27.9189 secs



100

```

QY 401 PGEPTAGSVGGSGGLALALALGGLGTA 442
DB 401 PGEPTAGSVGGSGGLALALALGGLGTA 442

RESULT 2
US 09-872 185B 2
Sequence 2: Application US/08/755,245
Patent No. US2004009423A1
GENERAL INFORMATION:
APPLICANT: Stern, David M.
APPLICANT: Herold, Kevin
APPLICANT: Yan, Shi Du
APPLICANT: Schmidt, Ann Marie
APPLICANT: Lanster, Ira
TITLE OF INVENTION: METHODS FOR TREATING INFLAMMATION
FILE REFERENCE: 0575/50159
CURRENT FILING DATE: 2001 06 01
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patent in version 3.1
SEQ ID NO 2
LENGTH: 442
TYPE: PRI
ORGANISM: Human
US 09-872 185B 2

Query Match
Best Local Similarity 100.0%; Score 1746; DB 18; Length 332;
Matches 332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AONTARIGEPVILKCKGAKKPPQPELWKNLIGRLTAEKVKVSPQGGQWDSVARVPLNG 60
DB 1 AONTARIGEPVILKCKGAKKPPQPELWKNLIGRLTAEKVKVSPQGGQWDSVARVPLNG 60

QY 61 SLFLPVGVLQDEGIFRCQAMNENKFKSNYPVPPVYQIPKPFIVDSASLTIAGVNVKVG 120
DB 61 SLFLPVGVLQDEGIFRCQAMNENKFKSNYPVPPVYQIPKPFIVDSASLTIAGVNVKVG 120

QY 61 SLFLPVGVLQDEGIFRCQAMNENKFKSNYPVPPVYQIPKPFIVDSASLTIAGVNVKVG 120
DB 61 SLFLPVGVLQDEGIFRCQAMNENKFKSNYPVPPVYQIPKPFIVDSASLTIAGVNVKVG 120

QY 121 TCVSSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 180
DB 121 TCVSSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 180

QY 181 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 240
DB 181 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 240

QY 241 AQPSPQIHMKKDVPFLPPESEVLILPEIGPQDQSYSCVATSSRSGPQESRAVSTISIE 400
DB 241 AQPSPQIHMKKDVPFLPPESEVLILPEIGPQDQSYSCVATSSRSGPQESRAVSTISIE 400

QY 401 PGEPTAGSVGGSGGLALALALGGLGTA 442
DB 401 PGEPTAGSVGGSGGLALALALGGLGTA 442

RESULT 4
US 08 755 245 2
Sequence 2: Application US/08/755,245
Patent No. US2004009423A1
GENERAL INFORMATION:
APPLICANT: Stern, David M.
APPLICANT: Schmidt, Ann Marie
APPLICANT: Wu, Jun
TITLE OF INVENTION: METHOD FOR TREATING SYMPTOMS OF DIABETES
FILE REFERENCE: 0575/50159
CURRENT FILING DATE: 1996 11 22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent in version 3.1
SEQ ID NO 2
LENGTH: 416
TYPE: PRI
ORGANISM: Bovine
US 08 755 245-2

Query Match
Best Local Similarity 82.8%; Score 1463; DB 1; Length 416;
Matches 278; Conservative 21; Mismatches 40; Indels 12; Gaps 2;

QY 2 ONITARIGEPVILKCKGAKKPPQPELWKNLIGRLTAEKVKVSPQGGQWDSVARVPLNG 61
DB 2 ONITARIGEPVILKCKGAKKPPQPELWKNLIGRLTAEKVKVSPQGGQWDSVARVPLNG 61

QY 62 LFLPVGVLQDEGIFRCQAMNENKFKSNYPVPPVYQIPKPFIVDSASLTIAGVNVKVT 121
DB 62 LFLPVGVLQDEGIFRCQAMNENKFKSNYPVPPVYQIPKPFIVDSASLTIAGVNVKVT 121

QY 122 TVSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 181
DB 122 TVSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 181

QY 14 TVSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 202
DB 14 TVSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 202

QY 182 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 240
DB 182 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 240

QY 203 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 262
DB 203 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 262

```

```

QY 1 AONTARIGEPVILKCKGAKKPPQPELWKNLIGRLTAEKVKVSPQGGQWDSVARVPLNG 60
DB 1 AONTARIGEPVILKCKGAKKPPQPELWKNLIGRLTAEKVKVSPQGGQWDSVARVPLNG 60

QY 23 AONTARIGEPVILKCKGAKKPPQPELWKNLIGRLTAEKVKVSPQGGQWDSVARVPLNG 82
DB 23 AONTARIGEPVILKCKGAKKPPQPELWKNLIGRLTAEKVKVSPQGGQWDSVARVPLNG 82

QY 61 SLFLPVGVLQDEGIFRCQAMNENKFKSNYPVPPVYQIPKPFIVDSASLTIAGVNVKVG 120
DB 61 SLFLPVGVLQDEGIFRCQAMNENKFKSNYPVPPVYQIPKPFIVDSASLTIAGVNVKVG 120

QY 83 SLFLPVGVLQDEGIFRCQAMNENKFKSNYPVPPVYQIPKPFIVDSASLTIAGVNVKVG 142
DB 83 SLFLPVGVLQDEGIFRCQAMNENKFKSNYPVPPVYQIPKPFIVDSASLTIAGVNVKVG 142

QY 121 TCVSSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 180
DB 121 TCVSSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 180

QY 14 TVSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 202
DB 14 TVSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 202

QY 181 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 240
DB 181 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 240

QY 203 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 262
DB 203 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 262

QY 241 AQPSPQIHMKKDVPFLPPESEVLILPEIGPQDQSYSCVATSSRSGPQESRAVSTISIE 299
DB 241 AQPSPQIHMKKDVPFLPPESEVLILPEIGPQDQSYSCVATSSRSGPQESRAVSTISIE 299

QY 263 AQPSPQIHMKKDVPFLPPESEVLILPEIGPQDQSYSCVATSSRSGPQESRAVSTISIE 322
DB 263 AQPSPQIHMKKDVPFLPPESEVLILPEIGPQDQSYSCVATSSRSGPQESRAVSTISIE 322

QY 300 EDCGEGTAGSVGGSGGLALALALGGLGTA 442
DB 300 EDCGEGTAGSVGGSGGLALALALGGLGTA 442

RESULT 4
US 08 755 245 2
Sequence 2: Application US/08/755,245
Patent No. US2004009423A1
GENERAL INFORMATION:
APPLICANT: Stern, David M.
APPLICANT: Schmidt, Ann Marie
APPLICANT: Wu, Jun
TITLE OF INVENTION: METHOD FOR TREATING SYMPTOMS OF DIABETES
FILE REFERENCE: 0575/50159
CURRENT FILING DATE: 1996 11 22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent in version 3.1
SEQ ID NO 2
LENGTH: 416
TYPE: PRI
ORGANISM: Bovine
US 08 755 245-2

Query Match
Best Local Similarity 82.8%; Score 1463; DB 1; Length 416;
Matches 278; Conservative 21; Mismatches 40; Indels 12; Gaps 2;

QY 2 ONITARIGEPVILKCKGAKKPPQPELWKNLIGRLTAEKVKVSPQGGQWDSVARVPLNG 61
DB 2 ONITARIGEPVILKCKGAKKPPQPELWKNLIGRLTAEKVKVSPQGGQWDSVARVPLNG 61

QY 62 LFLPVGVLQDEGIFRCQAMNENKFKSNYPVPPVYQIPKPFIVDSASLTIAGVNVKVT 121
DB 62 LFLPVGVLQDEGIFRCQAMNENKFKSNYPVPPVYQIPKPFIVDSASLTIAGVNVKVT 121

QY 122 TVSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 181
DB 122 TVSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 181

QY 14 TVSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 202
DB 14 TVSESYPACTLSWHLLGCFPLVDPNKGVSVAPOIRRHPELGLTQSLMLVTPARGGDP 202

QY 182 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 240
DB 182 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 240

QY 203 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 262
DB 203 RPTSCSFSPGLPRHRLAPLQPRVWVPLFVPLVPLVPLVPLVPLVPLVPLVPLVPLV 262

```





Best Local Similarity 25.9% Pred No 1 30-05  
Matches 79; Conservative 45; Mismatches 110; Indels 51; Gaps 12;

QY 4 TTAGGCEPLVLCCKGCAKPKPKKLEKNTQKFAKVLSPQGG-----PW 50  
Db 31 VNSAVGDTLILPKCK---LDVFNLMF---GK---WKYKPDGSPVFIARSTKKSVOY 80  
QY 51 LSV-----ARVLENSLIFAVVILQNGIFPQAMNPNKETSNY-----RVRVYQIP 99  
Db 81 DVMPEYKURKLNSENYYTLISNARISDEKKEVCMIV-----TEDNVFIAPTIVKVKQP 134  
QY 100 GKPEIVDSASELTAGVKNKGVTCVSEGSYPAGTILSHLDGKPLVPNEKGVSVKEGDTFRHP 159  
Db 135 SKPEIVSKALFETEQLKGLDGLSEUSYDGNITWYRNGKVLPLLEGAVVILFKEMDP 194  
QY 160 KTGIPFLLQSLVMTTPARGGDRPPTSCS---FSGNLPBRRALFTADIQHVRWE-IVVLEE 215  
Db 195 VTQLYTMTSTLEYKTK-ADIQMPTFCSTVYIGPS-----GQKIHSEQAVFDIYYPTQ 248  
QY 216 VCLVVEPEBMAVAPAGTIVTLCVEFAFSPQIHWKMGVVELFPLPSPVLLILPEIGPJDU 274  
Db 249 VTIOVLPPKNAIKEGDNITLKLONGNPPPEEFVLPQGEIHRSSNTIYTLMDVRNAT 308  
QY 275 GTYSC 279  
Db 309 GDYKC 313

RESULT 11  
US-09-808-602-70  
; Sequence 70, Application US/80808602  
; Publication No. US20020155115A1  
; GENERAL INFORMATION:  
; APPLICANT: Vernet, Corine A  
; APPLICANT: Fernandes, Ilma  
; APPLICANT: Shimkels, Richard A  
; APPLICANT: Heriman, John L  
; APPLICANT: Majumder, Kumud  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Mezes, Peter S  
; APPLICANT: MacDougall, John  
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 1866-647 CJP  
; CURRENT APPLICATION NUMBER: US/09/808,602  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 09/800,198  
; PRIOR FILING DATE: 2001-03-05  
; PRIOR APPLICATION NUMBER: 60/186,596  
; PRIOR FILING DATE: 2000-03-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 70  
; LENGTH: 793  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-808-602-70

Query Match 11.4% Score 201; DB 9; Length 793;  
Best Local Similarity 23.4% Pred No. 0.00022;  
Matches 97; Conservative 49; Mismatches 131; Indels 148; Gaps 19;  
QY 6 ARIGEPVLVLCCKGAPKPPORLEWKN-----TGRTEAMKVLSPQGGPMD 51  
Db 65 ANPQGVIVLCK-VEGTTPVQVSWRKNGAELPEGTHTTLLANGSLLIHHFRLEQGGSPD 123  
QY 52 -----SVAR----- 55  
Db 124 EGDYECVAQNRKFULLVSRKAKLQAATMSDFVHPQAVIGERGVAVRFOQIHLGPKPLIT 183  
QY 56 -----VLPNGSLFLPAVGIODEGIFRCQAMNRNKGKTSNVRV----- 95  
Db 184 WEKNRVPDITDDERYTLPLKGVLIQTIGLRAEDSGIFHCVASNIASVRVSHGARTVSSG 243  
QY 96 -----VLPNGSLFLPAVGIODEGIFRCQAMNRNKGKTSNVRV----- 95  
Db 184 WEKNRVPDITDDERYTLPLKGVLIQTIGLRAEDSGIFHCVASNIASVRVSHGARTVSSG 243

QY 96 ---YUIK-KKEIVDSASELTAGV-PNKVGTVCVSEGSYPAGTILSW-HLDGKPLVPNEKGV 150  
Db 244 SGTYK---EPTILVGNENLTLTTHOTAVLECVATGN-PRPIVSWSRIDGRPT--GVEGIO 297  
QY 151 VKROTRRHPDTGIFTASRLMVTIPAKGQIHPPPTSCSPSGCLPP-IPALPTAPIQPPVWER 210  
Db 298 VL-----GTGNLIISD---VIVQHSQ---VVVC---AANRRPTVRKTA---QGRLVVQ 339  
QY 211 VFLEEVOLVVEPEGGAAPGTTVTITCEVFAQSPQIHWKMGVPLPL-----PPSPV 263  
Db 340 APAEFVQ---HPOGISRPAGTTAMFTCOAGGEPHPVHTLKNCOVILCAGCHVRLKNNST 396  
QY 264 LILPEIGPOQGYTSCVATH-----SSHG-----POGSRVSVISIE 300  
Db 397 LSTISGVPEDEAIYQCAENIAGSSQASARLTLVWABGLPGPPRNVRVAVSSSTE 451

RESULT 12  
US-09-908-193-23  
; Sequence 23, Application US/0908193  
; Publication No. US20020192748A1  
; GENERAL INFORMATION:  
; APPLICANT: RASTELLI, LUCA  
; APPLICANT: SHIMKELS, RICHARD A.  
; APPLICANT: ZERHUSEN, BRYAN  
; APPLICANT: MALLYANKAR, URIEL M.  
; APPLICANT: PADIGARU, MURALIDHARA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY  
; FILE REFERENCE: 21102-062  
; CURRENT APPLICATION NUMBER: US/09/908,193  
; PRIOR FILING DATE: 2001-07-18  
; PRIOR APPLICATION NUMBER: 60/220,273  
; PRIOR FILING DATE: 2000-07-24  
; PRIOR APPLICATION NUMBER: 60/221,650  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: 60/221,233  
; PRIOR FILING DATE: 2000-07-27  
; PRIOR APPLICATION NUMBER: 60/220,912  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/218,875  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/218,870  
; PRIOR FILING DATE: 2000-07-18  
; PRIOR APPLICATION NUMBER: 60/218,901  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 793  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-908-193-23

Query Match 11.4% Score 201; DB 9; Length 793;  
Best Local Similarity 23.4% Pred No. 0.00022;  
Matches 97; Conservative 39; Mismatches 131; Indels 148; Gaps 19;  
QY 6 ARIGEPVLVLCCKGAPKPPORLEWKN-----TGRTEAMKVLSPQGGPMD 51  
Db 65 ANPQGVIVLCK-VEGTTPVQVSWRKNGAELPEGTHTTLLANGSLLIHHFRLEQGGSPD 123  
QY 52 -----SVAR----- 55  
Db 124 EGDYECVAQNRKFULLVSRKAKLQAATMSDFVHPQAVIGERGVAVRFOQIHLGPKPLIT 183  
QY 56 -----VLPNGSLFLPAVGIODEGIFRCQAMNRNKGKTSNVRV----- 95  
Db 184 WEKNRVPDITDDERYTLPLKGVLIQTIGLRAEDSGIFHCVASNIASVRVSHGARTVSSG 243  
QY 96 ---YQIPKPEIVDSASELTAGV-PNKVGTVCVSEGSYPAGTILSW-HLDGKPLVPNEKGV 150  
Db 244 SGTYK---EPTILVGNENLTLTTHOTAVLECVATGN-PRPIVSWSRIDGRPT--GVEGIO 297

```

QY 151 VKQTRRHHPEIGLFLQSELMVLPARGGDHPTFSCSFSGCLPRHRLRTAPQPRWEP 210
Db 298 VL -----GTGNLLISL VTVQHSQ ---VVV- AANPSTVVRETA QSELVQ 349
QY 211 VPLEEVLVVEPRAVAVPQGTVTLTCEVPAQSPQIHMKKGVPLPL-----PPSPV 263
Db 340 APAPFVQ---HPQISRPACTIAMPFTCAQCGEPHVTWLKNOGVLAGCHVRLKNNST 396
QY 264 LILPEIGPQAGTYSKVATH---SSBG-----PQESRAVSSISIE 300
Db 397 LSLSGVPEEAIVQVAENIAISSQASAKTLVLAENGLPGPPRNVRAVSSSTE 451

RESULT 14
US 09-800-198 59
: Sequence 59, Application US/0900108
: Publication No. US20040047816A1
: GENERAL INFORMATION:
: APPLICANT: Vetnet, Cornie AM
: APPLICANT: Fernandes, Edna
: APPLICANT: Shimkets, Richard A
: APPLICANT: Herrmann, John L
: APPLICANT: Majumder, Kumud
: APPLICANT: Mishra, Vishva
: APPLICANT: Mezes, Peter S
: APPLICANT: Rastelli, Luca
: TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: 15966-697
: CURRENT APPLICATION NUMBER: US/09/800,108
: CURRENT FILING DATE: 2001-03-05
: PRIOR APPLICATION NUMBER: 60/186,596
: PRIOR FILING DATE: 2000-03-04
: NUMBER OF SEQ ID NOS: 98
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 59
: LENGTH: 794
: TYPE: PRT
: ORGANISM: Mus musculus
US 09-800-198 59

Query Match 11.4%, Score 2017, DB 97, Length 794,
Best Local Similarity 23.4%, Pred. No. 0.00022,
Matches 97, Conservative 49, Mismatches 131, Indels 148, Gaps 19,

```

```

QY 6 ARIGEPVLVLCVAPKPKPPQPLEWFLN-----TQTEAMKVLSPQSGGWD 51
Db 65 ANGGQPIVLGCK VETTFPVQVSWPRKNQAPLPSTHTTILANGSTTHHPRFPGVSPSD 124
QY 52 -----SVAR ----- 55
Db 124 EGDYEVQNRKPLGLVSRKARKLAATMSDFHVPQAVTGFPGVAVPTVQTHLPLFLIT 184
QY 56 VLENGSLFLPAVTLDEGEPVLAAMN KETASNYEVV 92
Db 184 WEKNRPHIDDERYTLTKGVLIQTGLRAEISCFPHVASNIASVVSBGAKLIVSGSG 243
QY 96 YQIPGKPEIVASSELTAQV PKKVGTCVSEGSYPASTLSW ILAKKPLVPKRGVS 150
Db 244 SGTYK--EPTILVAPENITIVQTAVLEVATUN PPFIVSWLELGGFT GVSEIQ 297
QY 151 VKQTRRHHPTGLFLQSELMVLPARGGDHPTFSCSFSGCLPRHRLRTAPQPRWEP 210
Db 298 VL -----GTGNLLISL VTVQHSQ ---VVV- AANPSTVVRETA QSELVQ 349
QY 211 VPLEEVLVVEPRAVAVPQGTVTLTCEVPAQSPQIHMKKGVPLPL-----PPSPV 263
Db 340 APAPFVQ---HPQISRPACTIAMPFTCAQCGEPHVTWLKNOGVLAGCHVRLKNNST 396
QY 264 LILPEIGPQAGTYSKVATH---SSBG-----PQESRAVSSISIE 300
Db 397 LSLSGVPEEAIVQVAENIAISSQASAKTLVLAENGLPGPPRNVRAVSSSTE 451

```

```

RESULT 14
US 09-908-194-19
: Sequence 19, Application US/090908194
: Publication No. US20020619274A1
: GENERAL INFORMATION:
: APPLICANT: RASTELLI, LUCA
: APPLICANT: SHIMKETS, RICHARD A.
: APPLICANT: ZERHUSEN, RYAN
: APPLICANT: MALYANKAR, URIEL M.
: APPLICANT: PADIGARU, MORALITHARA
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
: FILE REFERENCE: 21402-062
: CURRENT APPLICATION NUMBER: US/09/908,194
: CURRENT FILING DATE: 2001-07-18
: PRIOR APPLICATION NUMBER: 60/222,274
: PRIOR FILING DATE: 2000-07-24
: PRIOR APPLICATION NUMBER: 60/221,650
: PRIOR FILING DATE: 2000-07-28
: PRIOR APPLICATION NUMBER: 60/221,234
: PRIOR FILING DATE: 2000-07-27
: PRIOR APPLICATION NUMBER: 60/220,912
: PRIOR FILING DATE: 2000-07-26
: PRIOR APPLICATION NUMBER: 60/218,875
: PRIOR FILING DATE: 2000-07-18
: PRIOR APPLICATION NUMBER: 60/218,870
: PRIOR FILING DATE: 2000-07-18
: PRIOR APPLICATION NUMBER: 60/218,901
: PRIOR FILING DATE: 2000-07-18
: NUMBER OF SEQ ID NOS: 74
: SOFTWARE: Patent In Ver. 2.1
: SEQ ID NO 19
: LENGTH: 1252
: TYPE: PRT
: ORGANISM: Mus musculus
US 09-908-194-19

Query Match 10.9%, Score 1923, DB 97, Length 1252,
Best Local Similarity 26.4%, Pred. No. 0.0011,
Matches 73, Conservative 39, Mismatches 120, Indels 45, Gaps 11,

```

```

QY 49 FWDGVARVLENSLFLPAVTLQDESLFPPVLAAMN PKKIPASNYEVP VTQIPEEP 102
Db 183 PERSGLTTPNGVQVLIIVQUSAGSTRGVASRQFASLIVALKRSLATKQD 242
QY 103 EIVDSASELTAIVPKR VTTVVSRSYPASTLSW ILKGF LVPEKGVSVKQETRRHP 159
Db 243 VVIVAAIENTVWSGCVWMEVAAAG PTFVSVWVQDKKI 284
QY 160 ETCLETH QSPMLVTPARGGDHPTFSCSFSGCLPRHRLRTAPQPRWEP 218
Db 285 ETDVVGRTPELLASAG PHEGVVYVHARKERTGRTAIAAARLRLVLAALISA 349
QY 275 VVEEKAVAFGTVTLTCEVPAQSPQIHMKKGVPLPLSHV ILPEIG 270
Db 340 PEALSKTPASTARVCHASGDFKALHWHASLPL KNGHVKVQGLASVITQIG 495
QY 271 POGDGTSCVATHSHSGPQESRAVSSISIEEGEPT 307
Db 496 TADAGTYQVAENGAGTAIAAALFVVVVRKGLPSAPT 443

```

```

RESULT 15
US 09-908-194-20
: Sequence 20, Application US/090908194
: Publication No. US2002019274A1
: GENERAL INFORMATION:
: APPLICANT: RASTELLI, LUCA
: APPLICANT: SHIMKETS, RICHARD A.
: APPLICANT: ZERHUSEN, RYAN
: APPLICANT: MALYANKAR, URIEL M.
: APPLICANT: PADIGARU, MORALITHARA
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
: FILE REFERENCE: 21402-062

```



```

; CURRENT APPLICATION NUMBER: US/09/908,193
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: 60/220,273
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: 60/221,650
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/221,233
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 60/220,912
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/218,875
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,870
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/218,901
; PRIOR FILING DATE: 2000-07-18
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 20
; LENGTH: 1253
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-908-193-20

Query Match      10.9%   Score 192.5, DB 9, Length 1253,
Best Local Similarity 26.4%   Pred. NO. 0.0011:
Matches 73; Conservative 49; Mismatches 120; Indels 45; Gaps 11;

QY 49 PWDVAVLPNGSLFLPAVLIQDENTFRQAMN----PDKETKSNYPV--VYQIPGKP 102
DB 183 PFSPIIIINWVQIHHVQSTAGSYPCVATNSAPQPSQFASIVALEGSLFATPQQD 242
QY 103 EIVDSASELTAGVPNK--VITVSEGSYPAGTLSW-HLDGKPLVPNEKGSVYKEQTRRH 159
DB 243 WIVAAPENTIVVSGSVVMEKVASAD-ETPEVSVVRQDKPI----- 284
QY 160 ETGIFTI-OSGLMVTARGGDPRPTSCSFSPGLPHHRAIRTAIPQRYWHPVPLFEVOL 218
DB 285 STEVIVLPTNLIILASQ----PHISGVYVCRANKPTRENTAAELPVLAAFAISQA-- 339
QY 219 WPEEGAVAPGSITVTLICVEVPAQPSQIHWKMDGVLPPLPSPV-----LILPEIG 270
DB 340 ---PEALSRTRASTARFVCKASCEPRPAHLHLHDGIPL--RPNGRVKKVQGGGSLVITQIG 395
QY 271 PQDQGYSCVATHSHGPOESRAVSISITPEPGEGET 307
DB 396 LODAGYYQCVARENSAGTACAAALAVVVRGLPSAPT 432

```

Search completed: May 30, 2003, 15:59:08  
Job time : 28.1712 secs



Genome version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 15:40:16 ; Search time 38.1351 seconds  
(without alignments)  
836.935 million cell updates/sec

Title: US-09-872-185b-2

Perfect score: 1766

Sequence: 1 AONTARIGEPVLKCKGAP.....GSGLGTLALAIILGLATA 332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 28324 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 28324

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR-73: \*  
1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	1766	100.0	404	161506	advanced glycosyla
2	1463	82.8	416	142879	advanced glycosyla
3	1381	78.2	402	70962	probable advanced
4	219.5	12.4	583	139428	alcam human
5	217.5	12.3	645	198240	cell surface glyco
6	217.5	12.3	1344	114316	right protein mo
7	212.5	12.0	1523	142478	acidic glycoprote
8	209	11.8	1612	730805	cell adhesion prot
9	208	11.8	1833	519247	adhesion molecule
10	207.5	11.7	588	145254	transmembrane rec
11	205.5	11.6	588	240506	perlecan precursor
12	203	11.5	1651	714360	glycocalyx
13	199.5	11.3	587	240464	glycocalyx
14	198	11.2	4391	138096	glycocalyx
15	194.5	11.0	584	150419	glycocalyx
16	190.5	10.8	702	136310	glycocalyx
17	189.5	10.7	423	729540	glycocalyx
18	189.5	10.7	1273	742406	glycocalyx
19	185.5	10.5	1521	834338	glycocalyx
20	185	10.5	3707	830252	glycocalyx
21	182	10.3	5175	720992	glycocalyx
22	182	10.3	6104	743200	glycocalyx
23	180.5	10.2	4569	143767	glycocalyx
24	179.5	10.2	4569	741009	glycocalyx
25	177	10.0	6642	729457	glycocalyx
26	175.5	9.9	344	156551	glycocalyx
27	172.5	9.8	739	150476	glycocalyx
28	169.5	9.6	458	140481	glycocalyx
29	169.5	9.6	518	744034	glycocalyx

## ALIGNMENTS

Result 1

161506

advanced glycosylation end-products receptor precursor - human

N: Alternate names: advanced glycosylation end-product-binding protein, 35K; glycoprotein  
C: Species: homo sapiens (man)  
C: Date: 24-May-1996 #sequence\_revision 07-feb-1997 #text\_change 16-Jul-1999

C: Accession: 161596; B42879; S27968  
R: Sugaya, K.; Fukagawa, T.; Matsumoto, K.; Mita, K.; Takahashi, E.; Ando, A.; Inoko, G.  
Genomics 23, 408-419, 1994

A: Title: Three genes in the human MHC class III region near the junction with the class II region of mouse mammary tumor gene int-3.

A: Reference number: A55562, MUID: 95137587, PMID: 7835890

A: Accession: 161596  
A: Status: nucleic acid sequence not shown; translation not shown; translated from GR/

A: Molecule type: DNA  
A: Residues: 1-404 <RES>

A: Cross-references: GR: D28760; NID: 9561657; FTDN: DAA05958.1; PID: g561659  
R: Neuber, M.; Schmidt, A.M.; Hroff, J.; Yan, S.D.; Wang, F.; Pan, Y.C.; Elliston, K.  
J. Biol. Chem. 267, 14998-15004, 1992

A: Title: Cloning and expression of a cell surface receptor for advanced glycosylation

A: Reference number: A42879; MUID: 92340547, PMID: 1378843

A: Accession: B42879  
A: Molecule type: mRNA

A: Residues: 1-299, R' 101-404 <RES>  
A: Cross-references: EMBL: M1211, NID: g190845; PID: AAA03574.1; PID: g190846

A: Experimental source: lung  
A: Note: sequence extracted from NCBI backbone (NCBIP: 109438)

A: Comment: Advanced glycosylation end-products are heterogeneous nonenzymatically glycosylated cellular function, thus contributing to tissue lesions in diabetes.

C: Comment: This receptor appears also to mediate the effects of amyloid beta peptide sites in the neurotoxic pathway that produces dementia in Alzheimer's disease.

C: Genetics:

A: Gene: GPR-AGER  
A: Cross-references: GDB: 306354; OMIM: 600214

A: Map position: 6p21.3  
A: Intron: 18/1, 19/1, 119/1, 140/3, 170/2, 231/1, 274/3, 322/2, 331/1, 373/2

C: Function:

A: Description: neuronal receptor for angiotensin, a G-protein-coupled protein involved in the

C: Keywords: Alzheimer's disease, glycoprotein, receptor, immunoglobulin homology

E: 1/2/Domain: signal sequence glycosylation end-product receptor, transmembrane protein

E: 2/3/4/Domain: extracellular status predicted <SIG>

E: 3/1/1/Domain: immunoglobulin homology <IM>

E: 137/2/Domain: immunoglobulin homology <IM>

E: 252/3/Domain: immunoglobulin homology <IM>

E: 345/362/Domain: transmembrane status predicted <TM>

E: 363/404/Domain: intracellular status predicted <IN>

E: 25/81/Binding site: carbohydrate (ASO) (covalent) status predicted

Query Match: 100.0%, Score 1766; DB 1: Length 404;

Query Match 82.8%; Score 1464; DB 1; Length 416;  
 Best Local Similarity 81.5%; Pred. No. 7.4e+90;  
 Matches 294; Conserved 176; Mismatches 40; Indels 12; Gaps 2;

QY 1 AGNITARIGEDVLKCGAPKPPORLEWKLNTGRTAKKVLSPQAGCPWPSVAPVPLNG 60  
 DB 23 AGNITARIGEDVLKCGAPKPPORLEWKLNTGRTAKKVLSPQAGCPWPSVAPVPLNG 82  
 QY 61 SLFLPVGICDDEGIFRQAMNPNKPKSNYPVAVYQIPCKPELVDSASELTACVKNKG 120  
 DB 83 SLFLPVGICDDEGIFRQAMNPNKPKSNYPVAVYQIPCKPELVDSASELTACVKNKG 142  
 QY 121 TCVSGSYVAGTISWHLCKKPLVNEKCVSKVKGTRRHPTGELTQSELMTVPARGGDP 180  
 DB 143 TCVSGSYVAGTISWHLCKKPLVNEKCVSKVKGTRRHPTGELTQSELMTVPARGGDP 202  
 QY 181 RPTSCSPSPGLPRRIALPTAPIQPVWPEVPLEEVLVVEPEGCAVAPGCTVTLICEVP 240  
 DB 203 RPTSCSPSPGLPRRIALPTAPIQPVWPEVPLEEVLVVEPEGCAVAPGCTVTLICEVP 262  
 QY 241 AQPSPQIHWKDKGVLLPSPVLLDPEISVLLFELGQDQCTYSCVATHSSHGQPSRAVSISIE 300  
 DB 263 AQPSPQIHWKDKGVLLPSPVLLDPEISVLLFELGQDQCTYSCVATHSSHGQPSRAVSISIE 322  
 QY 301 PGRHPTAGSGVGSGLGHALALGGLGCTA 332  
 DB 423 PGRHPTAGSGVGSGLGHALALGGLGCTA 354

RESULT 2  
 A42879  
 advanced glycosylation end-products receptor precursor - bovine  
 N:Alternate names: advanced glycosylation end product binding protein, 35k; glycoprotein  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 04 Mar-1993 #sequence revision 07 Feb-1997 #text\_change 16 Jul-1999  
 R:Accession: A42879; A42878; S27949  
 R:Receptor, M.; Schmidt, A.M.; Bretz, J.; Yan, S.D.; Wang, F.; Pan, Y.C.; Elliston, K.; St  
 J. Biol. Chem. 267, 14998-15004, 1992  
 A:Title: Cloning and expression of a cell surface receptor for advanced glycosylation en  
 A:Reference number: A42879; MUID:92440547; PMID:1378843  
 A:Accession: A42879  
 A:Molecule type: mRNA  
 A:Residues: 1-416 <NE>  
 A:Cross-references: GH-M91212; NID:q163650; PDB:AAA03575.1; PDB:1g163651  
 A:Experimental source: Lung  
 A:Note: sequence extracted from NBI backbone (NCBI:P109436)  
 A:Note: parts of this sequence, including the amino end of the mature protein, were dete  
 R. Schmidt, A.M.; Viana, M.; Gerlach, M.; Bretz, J.; Ryan, J.; Kao, J.; Esposito, C.; He  
 J. Biol. Chem. 267, 14987-14997, 1992  
 A:Title: Isolation and characterization of two binding proteins for advanced glycosylati  
 A:Reference number: A42878; MUID:92440546; PMID:137882  
 A:Accession: A42878  
 A:Molecule type: protein  
 A:Residues: 23-24, 'X', 26-37, 'X', 49-49, 'XX', 52-54 -SCH-  
 A:Experimental source: endothelial cells  
 A:Note: sequence extracted from NBI backbone (NCBI:P109434)  
 C:Comment: Advanced glycosylation end products are heterocyclics acetylenically glycos  
 cellular function, thus contributing to tissue lesions in diabetes.  
 C:Comment: This receptor appears also to mediate the effects of amyloid beta peptide on  
 ates in the neurotoxic pathway that produces dementia in Alzheimer's disease.  
 C:Function:  
 C:Description: neuronal receptor for amphotericin, a DNA-binding protein involved in neur  
 C:Superfamily: advanced glycosylation end products receptor; immunoglobulin homology  
 C:Keywords: Alzheimer's disease; glycoprotein; receptor; transmembrane protein  
 F:1-2/Domain: signal sequence #status predicted <SIG>  
 F:23-416/Product: advanced glycosylation end-products receptor PAGE #status predicted <M  
 F:23-354/Domain: extracellular #status predicted <EXT>  
 F:31-100/Domain: immunoglobulin homology <IM1>  
 F:136-209/Domain: immunoglobulin homology <IM2>  
 F:242-313/Domain: immunoglobulin homology <IM3>  
 F:355-372/Domain: transmembrane #status predicted <TM>  
 F:373-416/Domain: intracellular #status predicted <INT>  
 F:25,80/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:38-96,143-207,269-311/Disulfide bonds: #status predicted

Query Match 82.8%; Score 1464; DB 1; Length 416;  
 Best Local Similarity 81.5%; Pred. No. 7.4e+90;  
 Matches 294; Conserved 176; Mismatches 40; Indels 12; Gaps 2;

QY 2 QNITARIGEDVLKCGAPKPPORLEWKLNTGRTAKKVLSPQAGCPWPSVAPVPLNG 61  
 DB 24 QNITARIGEDVLKCGAPKPPORLEWKLNTGRTAKKVLSPQAGCPWPSVAPVPLNG 82  
 QY 62 LFLPVGICDDEGIFRQAMNPNKPKSNYPVAVYQIPCKPELVDSASELTACVKNKG 121  
 DB 83 LFLPVGICDDEGIFRQAMNPNKPKSNYPVAVYQIPCKPELVDSASELTACVKNKG 142  
 QY 122 CVSGSYVAGTISWHLCKKPLVNEKCVSKVKGTRRHPTGELTQSELMTVPARGGDP 181  
 DB 143 CVSGSYVAGTISWHLCKKPLVNEKCVSKVKGTRRHPTGELTQSELMTVPARGGDP 202  
 QY 182 PTFSCSPSPGLPRRIALPTAPIQPVWPEVPLEEVLVVEPEGCAVAPG 240  
 DB 203 PTFSCSPSPGLPRRIALPTAPIQPVWPEVPLEEVLVVEPEGCAVAPG 262  
 QY 241 GTVTLCEVAGTISWHLCKKPLVNEKCVSKVKGTRRHPTGELTQSELMTVPARGGDP 290  
 DB 263 GTVTLCEVAGTISWHLCKKPLVNEKCVSKVKGTRRHPTGELTQSELMTVPARGGDP 322  
 QY 291 SKAVSISIEPEGCPHPTAGSGVGSGLGHALALGGLGCTA 331  
 DB 323 SKAVSISIEPEGCPHPTAGSGVGSGLGHALALGGLGCTA 354

RESULT 3  
 I09062  
 probable advanced glycosylation end-products receptor precursor mouse  
 N:Alternate names: RAGE  
 C:Species: Mus musculus (house mouse)  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21 Jan 2000  
 C:Accession: I09062  
 R:Kowch, S.; Mahaffey, G.; Glin, S.; Abcarin, M.E.; Bankers, C.; Lasky, S.; Koretz, C.;  
 submitted to the EMBL Data Library, October 1997  
 A:Description: Sequence of the mouse major histocompatibility locus class III region.  
 A:Reference number: Z16543  
 A:Accession: I09062  
 A:Status: preliminary, translated from Gh/Embl/I09062  
 A:Molecule type: DNA  
 A:Residues: 1-402 <RW>  
 A:Cross-references: EMBL:AF040001; NID:q256445; PDB:q254950  
 C:Genetics:  
 A:Gene: RAGE  
 A:Map position: 17  
 A:Notes: 24, 'X', 26-37, 'X', 49-49, 'XX', 52-54 -SCH-  
 C:Superfamily: advanced glycosylation end products receptor; immunoglobulin homology  
 C:Keywords: receptor; transmembrane protein  
 F:1-100/Domain: immunoglobulin homology <IM>

Query Match 78.2%; Score 1381; DB 2; Length 402;  
 Best Local Similarity 79.9%; Pred. No. 2e+84;  
 Matches 263; Conservative 19; Mismatches 45; Indels 2; Gaps 2;

QY 2 QNITARIGEDVLKCGAPKPPORLEWKLNTGRTAKKVLSPQAGCPWPSVAPVPLNG 61  
 DB 24 QNITARIGEDVLKCGAPKPPORLEWKLNTGRTAKKVLSPQAGCPWPSVAPVPLNG 82  
 QY 62 LFLPVGICDDEGIFRQAMNPNKPKSNYPVAVYQIPCKPELVDSASELTACVKNKG 121  
 DB 83 LFLPVGICDDEGIFRQAMNPNKPKSNYPVAVYQIPCKPELVDSASELTACVKNKG 142  
 QY 122 CVSGSYVAGTISWHLCKKPLVNEKCVSKVKGTRRHPTGELTQSELMTVPARGGDP 181  
 DB 143 CVSGSYVAGTISWHLCKKPLVNEKCVSKVKGTRRHPTGELTQSELMTVPARGGDP 201  
 QY 182 PTFSCSPSPGLPRRIALPTAPIQPVWPEVPLEEVLVVEPEGCAVAPG 241  
 DB 202 PTFSCSPSPGLPRRIALPTAPIQPVWPEVPLEEVLVVEPEGCAVAPG 261

QY 242 QPSTQTHMKWGV...  
 DB 262 QPPQVHWKDCGAPLAFSLVLLPEVGEDEGYSVATHPSHGQSFSPVSTVET 321  
 QY 302 GREGPIAGSVGSGCTGTAALACILGCG 330  
 DB 322 GDCGPAFGSVGSGCTGTAALACILGCG 350

## RESULT 4

139428  
 A:Accession: 139428  
 C:Species: Homo sapiens (man)  
 C:Date: 01-Sep-1996 #sequence\_revision 16-Sep-1996 #text\_change 21-Jul-2000  
 P:Brown, M.A.; Patel, D.D.; Li, X.; Modrell, B.; Malacko, A.P.; Wang, W.C.; Marquardt, H.  
 J. Exp. Med. 181, 2213-2220, 1995  
 A:Title: Cloning, mapping, and characterization of activated leukocyte-cell adhesion mol  
 A:Reference number: 139428, PMID: 9527947, PMID: 7762007  
 A:Accession: 139428  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: mRNA  
 A:Residues: 1-583 <RES>  
 A:Cross-references: GB:138008, NID:9486257, FIIN:AA659499.1, PID:9486258

Query Match 12.4%, Score 219.5, DB 2, Length 583  
 Best Local Similarity 25.9%, Pred. No. 3.2e 07  
 Matches 79, Conservative 45, Mismatches 130, Indels 51, Gaps 12

QY 4 ITARIGEPVAVLCKGAPKPKPPORLEWKLNTGTEAWKVLSPGGG-----PW 50  
 DB 31 VNSAYGDTIIIPCF-----LVPGLNMF-----GK--WKYKPGSPVFTAFRSTKKSQY 80  
 QY 51 DSV-----ARVLPNSGLPVGIDGIFRCQAMNNGKETKSY-----PVRYQTP 99  
 DB 81 DIWPKYKDRINSENVTLSNARSDEKRFVCMIV-----TEDVFPAPITIVKVPKOP 134  
 QY 100 GKPEIVSASELTAVNFKVGVGVSEGSYPARTLSMILDKPLVPNEKGVSEKTRRRP 159  
 DB 135 SKPEIVSKALFLETKLKGDCISDSYDGNITWYRNGKVLHLEGAVVIFKEMDP 194  
 QY 160 ETGLFTIQSLFMTVPARGGPPPTSCS---ESPGLPPHRLTAIOPRWVF-PVPLEE 215  
 DB 195 VTQLYMTSTLEKTKK-ADIQPFCTSVTYGPS-----GKLIUSEAVAVEDVYPTEQ 248  
 QY 216 VQLVWPEPESAVAPAGTIVLTIEVPAIPSP-QLHMKWLVPLPSPVLLILPEIGPQD 274  
 DB 249 VTQVILPPNNAKESGNTLTKGNGNPPPEPELEYLPQCPGKSSNIYTLMDVKRNAT 308  
 QY 275 GTYSC 279  
 DB 309 GDYCK 313

## RESULT 5

138049  
 A:Accession: 138049  
 C:Species: Homo sapiens (man)  
 C:Date: 01-Mar-1996 #sequence\_revision 31-Mar-1996 #text\_change 21-Jul-2000  
 P:Seis, C.; Kirsch, K.; Pfeilbacher, H.; Pfeilbacher, G.; Johnson, T.P.  
 Proc. Natl. Acad. Sci. U.S.A. 93, 8514-8518, 1996  
 A:Title: Germ-line origin of the melanoma-associated glycoprotein, MUC18: implication  
 A:Reference number: 138049, PMID: 93391384, PMID: 8378324  
 A:Accession: 138049  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-646 <RES>  
 A:Cross-references: EMBL: X6024, NID: 943891, FIIN: AA48332.1, PID: 9425693  
 A:Note: this reference also contains a correction to A34507  
 R. Lehmann, J.M., R. Kiehlmann, G. Johnson, J.P.

Proc. Natl. Acad. Sci. U.S.A. 96, 9891-9895, 1999  
 A:Title: MUC18, a marker of tumor progression in human melanoma, shows sequence simil  
 A:Reference number: A34507, PMID: 9009368, PMID: 2602381  
 A:Accession: A34507  
 A:Molecule type: mRNA  
 A:Residues: 1-71, SSSVCARARANIGSTSK, 91-120, LGPRSTASSASTKLRM, 139-587, AAVQALREAG  
 A:Cross-references: GB:M29277, GB:M28882  
 A:Note: this sequence has been corrected in 138049  
 C:Genetics:  
 A:Gene: GDB:MCM; MUC18; CD146  
 A:Cross-references: GDB:304548, OMIM:155735  
 A:Introns: 23/1, 64/3, 134/1, 157/3, 187/1, 247/1, 287/3, 342/1, 381/3, 429/1, 469/3;  
 C:Keywords: glycoprotein, transmembrane protein

Query Match 12.3%, Score 217.5, DB 2, Length 646  
 Best Local Similarity 23.3%, Pred. No. 5e-07  
 Matches 95, Conservative 55, Mismatches 135, Indels 123, Gaps 19

QY 4 IIAKGGPHVLKCKGAPKPP-ORLEW--KLNTGTEAWKVLSPG---GGPMDSVARVL 57  
 DB 36 VEVEGSLALDKSLSSSSVNLSHVDFSVIHKETLIFVPV--GSSFPPEYEQPLSLQ 95  
 QY 58 FNG-SLELPVAGVIGEGIFRCQAMNNGKETKS-NYRV--RVYVLPKPEI-----V 105  
 DB 96 IMAA LALQVTPQGERIFLQ---GKPPKSEYKRIQKVKYKAPPEEPNIQVNPGLGPV 150  
 QY 106 DSASHLTAGVPNKVGVGVSEGSYPAGTLSMILDKPLVPNEKGVSEKTRRRHPETGLFT 165  
 DB 151 NSKE-----PEEVAICGPGNYPLPQVIVYKNGKPLKEKNRVL-I-QSSUTVESSGLYT 203  
 QY 166 LOSELMTVPARGGPPPTSCSPGCLPPHRLALR-----TAPI---OPRW---EPVPL- 213  
 DB 204 LQSLIKAGLVK-EDKDAOFYCELYRLPSPGNHMKESREVIVPEYPTKVLVEVPVGM 262  
 QY 214 -----EE----- 215  
 DB 263 KGDHVEIRCLADGNIPPHPSIKONPSTREAREFTTNDGVLVLEPARKHSCRYEQCA 322  
 QY 216 -----VLVVEPEG-----GAVAF-----GGTIVLTCEVFAQSPQIHMMKDV 254  
 DB 323 WNLDMISLLSEPVELLVNVYSDVRVSPAAPERUGSSSLITLTCFAESSDLEFOWLREET 382  
 QY 255 PLPPLPPSVLLPEIGPQDGGIYSCVAT-HSSHGQKSKAVSISLIEP 301  
 DB 393 DQVLEKGVPLQHLDKREAGGKVCASVPSIPGLNRTULVVKLAIFGP 430

## RESULT 6

114316  
 A:Accession: 114316  
 C:Species: Mus musculus (house mouse)  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 R. Yuan, S.F.; Cox, J.A.; Basika, G.K.; Lee, P.Y.H.P.  
 Submitted to the EMBL Data Library, April 1998  
 A:Reference number: 217975  
 A:Accession: 114316  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1344 <YUA>  
 A:Cross-references: EMBL: AF06570, NID: 94206485, FIIN: AAD1428.1

Query Match 12.3%, Score 217, DB 2, Length 1344  
 Best Local Similarity 26.7%, Pred. No. 1.2e-06  
 Matches 85, Conservative 39, Mismatches 119, Indels 76, Gaps 15

QY 9 GPHVLKCKGAPKPPORLEWKLNTGTEAWKVLSPGGGPPMDSVAR-----VLPNG 60  
 DB 56 GEPATLPGR-AEGRPFPNIWYKNGARV-----ATAREDPPAHLRLPSG 99  
 QY 61 SLEFLAV-----GIGEGIFRCQAMNNGKETKS-NYRVPV-----VYIFGSKPEIVDSA 108  
 DB 100 ALFPFPIVHGKSHPLKGVYTCVAKNYLGAAASRNASLEAVLKLDPQSGCNVVV----- 155

QY 109 SELLACVINKVAVLVSAGSYFAGLISWRLLQKHLVNEKEGVSVKQIHRHHHGHILFIHQ 167  
 DB 156 AVGEP-AMVSVVFKRHP 111 111 111 111 111 111 111 111 111 111  
 QY 168 SELMVVDAGGDPRTFSYSPSSGILPRHRLRTAPLQPKWHPVPLVPLVQVIVPEGGA 226  
 DB 196 GGLKMSHIFKSD AVMYKCVAS NMAGKESCAA 111 111 111 111 111 111 111 111  
 QY 227 VAKSGVTLLCPVFAQESFQIHWKKGVPPLP 111 111 111 111 111 111 111 111  
 DB 247 VLADAVNPLCFVGLQDQNLHKKKDDGELPAQRYEIRSHSLMIDVSSSEDEGTYTCA 406  
 QY 282 TRSHSGDPQSRVAVSISL 298  
 DB 407 ENSVGRFAFASGSLSVHV 423

RESULT 7  
 150478  
 neurofilin - goldfish (fragment)  
 C:Species: Carassius auratus (goldfish)  
 C:Date: 14-Sep-1996 #sequence revision 14 Sep 1996 #next\_change 14 Sep 1996  
 C:Accession: 150478  
 Kilaessing, U.; Giordano, S.; Stecher, H.; Lottspeich, F.; Stuermer, C.A.  
 Differentiation 56, 21-29, 1994  
 A:Title: Molecular characterization of fish neurofilin, a growth associated cell surface protein DM-GRASP/sc-17BEN.  
 A:Reference number: 150478; MID:94299040; PMID:8026643  
 A:Accession: 150478  
 A:Status: preliminary; translated from GB/EMBL/Genbank  
 A:Molecule type: mRNA  
 A:Residues: 1523<LAE.  
 A:Cross-references: GB:L25056; NID:407418; PID:407419

Query Match 12.8% Score 212.5; DB 2; Length 523;  
 Best Local Similarity 25.5% Pred. No. 8 3e 07;  
 Matches 76; Conservative 40; Mismatches 133; Indels 49; Gaps 11;  
 QY 9 GEHLVLKCGAKKKRQPLEWINTGPTAEWVLSQGRGPMDSVAR 55  
 DB 2 GEHLVVKCGAKKKRQPLEWINTGPTAEWVLSQGRGPMDSVAR 52  
 QY 56 VLENGSLFPAVGLQDGLERCOAMNRNKKFLKSNYRVVYVGLCKKPELVISA 108  
 DB 53 GYKSRVSIANSLLIARGSLAQVFTQMVVSFTNLEYS-VEYVHKKESAPVIRKNA 111  
 QY 109 SELLACVINKVAVLVSAGSYFAGLISWRLLQKHLVNEKEGVSVKQIHRHHHGHILFIHQ 168  
 DB 112 KELENGKLTQIGRCVVENANPQADLLKKNNQTLIDUCKTLLTSLTLLKDKLIGLSSTSS 171  
 QY 169 ELMVTPARKGDPRTFSYSPSSGILPRHRLRTAPLQPKWHPVPLVPLVQVIVPEGGA 222  
 DB 172 RQYV-ARKGWSQPTCA AKH VNGPQV VSPHSPHYPYTERKVSQVVS 220  
 QY 223 EGVAVAPGQTVLLIPEVIAQSPQLQHWKKGVPPLP 274  
 DB 221 Q SPTEPEDVTLKALQENPPTSPENIKKKVTVTLKLVYTLVATVRAVSGSLYK 277

RESULT 8  
 140805  
 dut11 protein mouse  
 N:Alternate names: Transmembrane receptor protein Rebol homolog  
 C:Species: Mus musculus (house mouse)  
 C:Date: 22 Oct 1999 #sequence revision 22 Oct 1999 #next\_change 22 Oct 1999  
 C:Accession: 140805  
 Kow, M.C.; Iow, N.; Fordham, R.; Rabbits, P.  
 submitted to the EMBL Data Library, July 1998  
 A:Description: The mouse homologue of human DUT11/H rebo1 gene; protein sequence and chr  
 A:Reference number: 220879  
 A:Accession: 140805  
 A:Status: preliminary; translated from GB/EMBL/Genbank

A:Molecule type: mRNA  
 A:Residues: 11612<WIM.  
 A:Cross-references: EMBL:Y17794; NID:q149714; PID:q149714; PIR:q149714; PIR:q149714  
 A:Experimental sources: Brain  
 C:Genetics:  
 A:Gene: dut11  
 A:Map position: 16

Query Match 11.8% Score 209; DB 2; Length 1612;  
 Best Local Similarity 25.4% Pred. No. 6 2e 06;  
 Matches 78; Conservative 44; Mismatches 143; Indels 52; Gaps 12;  
 QY 9 GEHLVLKCGAKKKRQPLEWINTGPTAEWVLSQGRGPMDSVAR 55  
 DB 43 GEPATLNCK AEGPTPTIEWKGGIEVETIKD DPR 94  
 QY 68 ----GTLQPLPRPCAMNPNKCFKSNYPVAVVYVPLPQFDELVSASLTAQVIN KYV 122  
 DB 95 HGRKSRPRPRVYICVARNYATGFAVSHNASLEVALH RDEPKJNSVMVAVGILFAVMEC 192  
 QY 124 VSGSYGAGTLSWILDGKPTVPNEKGVSVKQIHRHHHGHILFIHQ 182  
 DB 153 QPFGCHPEPTLSWKKDSPLQDKERITIR 195  
 QY 184 IFSCSFSGILPRHRLRTAPLQPKWHPVPLVPLVQVIVPEGGA 249  
 DB 196 ----AKYVGVGINMVGRESEVAELLVLERKSVKQIHRHHHGHILFIHQ 246  
 QY 240 PAQSPQIHWKKGVPPLP 294  
 DB 237 ECDVTVVAKKSGACELPRKRYLFGKHLTKKAVIAHMYLYTCA ENMVGRKASA 404  
 QY 295 SLSIIEP 401  
 DB 405 TLTQEP 411

RESULT 9  
 S19247  
 cell adhesion protein gp160 Drk fruit fly (Drosophila melanogaster)  
 C:Species: Drosophila melanogaster  
 C:Date: 14-Jan-1995 #sequence revision 14 Jan 1995 #next\_change 24 Sep 1999  
 C:Accession: S19247  
 Kaplido, D.; Campuzano, S.; Koda, T.; Modolell, J.; Barbacid, M.  
 EMBO J. 11, 491-404, 1992  
 A:Title: Drk, a Drosophila gene related to the tyk family of neurotrophin receptors.  
 A:Reference number: S19247; MID:92164624; PMID:1371458  
 A:Accession: S19247  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 11033<DOL.  
 A:Cross-references: EMBL:X63453; NID:q7883; PIR:q7883; PIR:q7883  
 C:Genetics:  
 A:Gene: Flybase:TK48b  
 A:Cross-references: Flybase:FB00004839  
 C:Species: Flybase:FB00004839  
 C:Keywords: ATP  
 E:53-526, kinase, protein kinase homology, KIN.  
 E:538-706/region: protein kinase ATP binding motif

Query Match 11.8% Score 208; DB 2; Length 1044;  
 Best Local Similarity 22.4% Pred. No. 3 6e 06;  
 Matches 88; Conservative 56; Mismatches 159; Indels 106; Gaps 18;  
 QY 12 EVLEK--KVAI KFFPQPLQWELNTLPTLQWVLSFQWVAVVAFHLSLFLFAV 68  
 DB 133 LALKCHVESASQDEDLFTWVNSKLSLTKWNVQ 179  
 QY 69 IOPGGLPRCAMNPNKCFKSNYPVAVVYVPLPQFDELVSASLTAQVIN KYV 192  
 DB 180 SEIDGLYPRATASNAAGPVMKSGYGVYSSVYVTLPLPPPKNPPMSEWIKQIFLPR 246  
 QY 194 ----ELVLCASLTLACVPRKVGTCVGRSYVACI 438

DB 237 KRCCAAAGLEALIAAPADIRIVQGFPGQGIKGEHFALTCIXKLPDELKAKNORLUKRWKKD 296  
 QY 139 GKPLVPNKGVS-----VKQTR--RHPETGLTLOSELMTVPAPGGDPP 181  
 DB 297 GKLLRQVFGSGAPFCHSHSKDALLBEDAKVLHKONG--TLSPASTIASDAG----- 350  
 QY 182 PTFSCSPSGIPRRHAPLTPATIQPVVRRVPLRFVOLVVERPEGAVANGGIVT LTCVPF 240  
 DB 351 QVQCQLQ--LEAHAPINSPPGLEV-----LEQKFVPOPTSKNLELDVAVKAVHCKAQ 402  
 QY 241 AQSPQIHMKKGVPLERF--PSPVILPEPQVQVQVSCVAVHSSHGTOESRAV 294  
 DB 403 GILPVPVQVVKUGENTTIPHVVEDANCTLFRNVNSHPGNYTCLATNSOQINATVAI 462  
 QY 295 SISIEBEGGPTAGSGGSLGTLALALGILG 327  
 DB 463 NV-VVTPKFSVPVPGPIRTSKQCTVVMHCQAIC 494

## RESULT 10

A45254

surface glycoprotein BEN precursor - chicken

C:Species: Gallus gallus (chicken)

C:Date: 27 Jun 1994 #sequence\_revision 27-Jun-1994 #text\_change 21-Jul-2000

C:Accession: A45254; S19202

R:Pourquie, O.; Corbel, C.; Le Caer, J. P.; Bossier, J.; Le Douarin, N.M.

A:Title: BEN, a surface glycoprotein of the immunoglobulin superfamily, is expressed in

Proc. Natl. Acad. Sci. U.S.A. 89, 5261-5265, 1992

A:Reference number: A45254; MUID:92302224; PMID:1608932

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-588 &lt;POU&gt;

A:Cross-references: EMBL:X64301; NID:363087; PIDN:CAA45579.1; FID:963088

C:Keywords: glycoprotein

Query Match 11.7%; Score 207.5; DB 2; Length 588;

Best local similarity 20.9%; Pred. No. 26-06;

Matches 84; Conservative 46; Mismatches 147; Indels 125; Gaps 11;

QY 4 ITARIQEPLVIAKCKGAPKPPQPLFWK--LNTGPTTEAKWVLSPQGG----- 48  
 DB 37 VNAVYGDITMPC-----RLEVPGGLMFGK---WKYMPNCPVFIAFRSSTKKNV 84  
 QY 49 PWDGV--ARVLPGSLFLPAVGIDQEGIFRCQANRNKGETKSNRYRVVYQIPGKP 102  
 DB 85 QYDVPDYKDRLSLSENYTLISKNARISDEKRFVCMIVTEDDVSEPTV-VKVFQPSQP 143  
 QY 103 EIVDSASELTAGVFNKVCVSGSPAGTLSWHLDGKPLVPNEKGVSKQTERHPETG 162  
 DB 144 EILHQADFLETEKLMGELVVRDSYPEGNVITWYKGRVQLQPFVEVVINLRKYENRSTG 203  
 QY 163 LFTLOSELMTVPARGCDPRPTFSCSPGLPFRHRLARTAPIQRPVWE--PVPLEEVQIWE 221  
 DB 204 LFTMTSSLOYMPTK-EDANAKFTCIIVTYHGPSQ--KTIQSEPVVDVHYVTEKVTIRVL 260  
 QY 222 PEGGAVAPGGTITICEVPAQSPQ----- 246  
 DB 261 SQSSTIKEGDNVILKCSGNGNPPQPFLEYIPGETEGIRSSDTVTMTDVRNATGEYKCS 320  
 QY 247 -----HMKDGVPLPL 258  
 DB 321 LIDKSNMDDTTITVHYLDLQTPSGEVTQIGEALPVSCITSSSRNATVFWIKDTRNKT 380  
 QY 259 PPSVILPEIPGQDQGVSCVATHSSHGPOESRAVSIITIF 300  
 DB 381 SPS-----FSSLOYQVAGNYICFTTHKVEGLKKRKLKLVIVE 418

## RESULT 11

JH0506

adhesion molecule SCI precursor - chicken

C:Species: Gallus gallus (chicken)  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 21-Jul-2000  
 C:Accession: JH0506; PS0270  
 R:Tanaka, H.; Matsui, T.; Arita, A.; Tomura, M.; Kubota, I.; McFarland, K.C.; Kohr, B.  
 Neuron 7, 535-545, 1991  
 A:Title: Molecular cloning and expression of a novel adhesion molecule, SCI.  
 A:Reference number: JH0506; MUID:92030150; PMID:1931049  
 A:Accession: JH0506  
 A:Molecule type: mRNA  
 A:Residues: 1-588 <TAN>  
 A:Cross-references: GB:S63376; NID:9238000; PIDN:AAR20170.1; FID:9238001  
 A:Experimental source: embryo  
 A:Accession: PS0270  
 A:Molecule type: protein  
 A:Residues: 34-48 <TAN1>  
 C:Comment: This protein is un-queely and transiently expressed on spinal cord motoneur  
 C:Keywords: glycoprotein; transmembrane protein  
 F:1-33/Domain: signal sequence #status predicted <SIG>  
 F:34-588/Product: adhesion molecule SCI #status predicted <ADH>  
 F:500-523/Domain: transmembrane #status predicted <TRA>  
 F:101-173/199,271,312,366,462,485,504/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 11.6%; Score 205.5; DB 2; Length 588;

Best local similarity 24.8%; Pred. No. 2.8e-06;

Matches 75; Conservative 46; Mismatches 135; Indels 47; Gaps 11;

QY 4 ITARIQEPLVIAKCKGAPKPPQPLFWK--LNTGPTTEAKWVLSPQGG----- 48  
 DB 37 VNAVYGDITMPC-----RLEVPGGLMFGK---WKYMPNCPVFIAFRSSTKKNV 84  
 QY 49 PWDGV--ARVLPGSLFLPAVGIDQEGIFRCQANRNKGETKSNRYRVVYQIPGKP 102  
 DB 85 QYDVPDYKDRLSLSENYTLISKNARISDEKRFVCMIVTEDDVSEPTV-VKVFQPSQP 143  
 QY 103 EIVDSASELTAGVFNKVCVSGSPAGTLSWHLDGKPLVPNEKGVSKQTERHPETG 162  
 DB 144 EILHQADFLETEKLMGELVVRDSYPEGNVITWYKGRVQLQPFVEVVINLRKYENRSTG 203  
 QY 163 LFTLOSELMTVPARGCDPRPTFSCSPGLPFRHRLARTAPIQRPVWE--PVPLEEVQIWE 221  
 DB 204 LFTMTSSLOYMPTK-EDANAKFTCIIVTYHGPSQ--KTIQSEPVVDVHYVTEKVTIRVL 260  
 QY 222 PEGGAVAPGGTITICEVPAQSPQ-----HMKDGVPLPLPPSPVILPEIPGQDGT 276  
 DB 261 SQSSTIKEGDNVILKCSGNGNPPQPFLEYIPGETEGIRSSDTVTMTDVRNATGE 316  
 QY 277 YSC 279  
 DB 317 YKC 319

## RESULT 12

T14160

transmembrane receptor protein Robol - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T14160

R:Kidd, T.; Brose, K.; Mitchell, K.J.; Fetter, R.D.; Tessier-Lavigne, M.; Goodman, C.

Cell 92, 205-215, 1998

A:Title: Roundabout controls axon crossing of the CNS midline and defines a novel sub

A:Reference number: Z17897; MUID:98117249; PMID:9458045

A:Accession: T14160

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1651 &lt;KID&gt;

A:Cross-references: EMBL:AF041082; NID:29111215; FID:g2811216; PIDN:AAC39960.1

C:Function:

A:Description: appears to function as the gatekeeper controlling midline crossing

C:Keywords: transmembrane protein

Query Match

Best local similarity 11.5%; Score 203; DB 2; Length 1651;

Matches 76; Conservative 46; Mismatches 133; Indels 52; Gaps 12;





Db 113 SISKVPRQDNARIFDQVGAUSQGV-----GSRTELYTKIHAPPEITFN-- 158  
 QY 110 ELTAIVPN KVAIVVVEESYFATILSWILFHKPLVNEKQVSVKKEQTPHPETGL 163  
 Db 159 --SAGIHALSNIMIKAVGISENFPSPNIIWYKCEPFIQFFPTKIIITLVPE-SNGL 215  
 QY 164 FIDASELMVITVAPGCHREKIFPSSSEFNGLIKHFAIKTAPIQPRVWEIV--FLREVLAVVE 221  
 Db 216 YIVVSTILFSKVIPE-EDENSLFPHATVIVWL--GQMPTEF-SERVVWTVFVETHEIVELVA 271  
 QY 222 IFDGAVALGGTVITCEVPAQFSGNCHMMK-----DCVGLFIPSPVLLIPEIQND 273  
 Db 272 INACTIVREGDULKIVLDADGNPAIVSEFRRKRGDSWQDMISLADINIGVLMHLNVS KSS 331  
 QY 274 QGTYSQ 279  
 Db 332 SGLVQC 337

Search completed May 30, 2003, 15:18:27  
 Job time : 40.1351 secs

F:123-133/Domain: I<DOM1>  
 F:194-530/Domain: II<DOM2>  
 F:190-234/Domain: IFL receptor binding repeat homology<ELL1>  
 F:295-319/Domain: ILL receptor binding repeat homology<ELL2>  
 F:325-329/Domain: ILL receptor binding repeat homology<ELL2>  
 F:368-403/Domain: ILL receptor binding repeat homology<ELL4>  
 F:531-676/Domain: III<DOM3>  
 F:1159-1266/Domain: laminin-type EGF-like homology<EGF>  
 F:1563-1610/Domain: laminin-type EGF-like homology<SP37>  
 F:1613-1668/Domain: laminin-type EGF-like homology<LE18>  
 F:1677-1686/Domain: IV<DOM4>  
 F:2507-2034/Domain: transmembrane status predicted<TM>  
 F:3687-4391/Domain: V<DOM5>  
 F:3845-3880/Domain: EGF homology<EGF>  
 F:3888-3921/Domain: EGF homology<EGF>  
 F:3983-4156/Domain: laminin 5 repeat homology<E2>  
 F:4147-4175/Domain: EGF homology<EGF2>  
 F:4149-4151/Domain: EGF homology attachment (I-P-E) motif  
 F:4299-4301/Domain: motor neuron attachment (I-P-E) motif  
 F:65-7176/Binding site: heparan sulfate (Ser) (covalent) #status predicted  
 F:89-444,1755,2141,4072,4135,4279,4789,9846,4968/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:2995,3933,4179/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

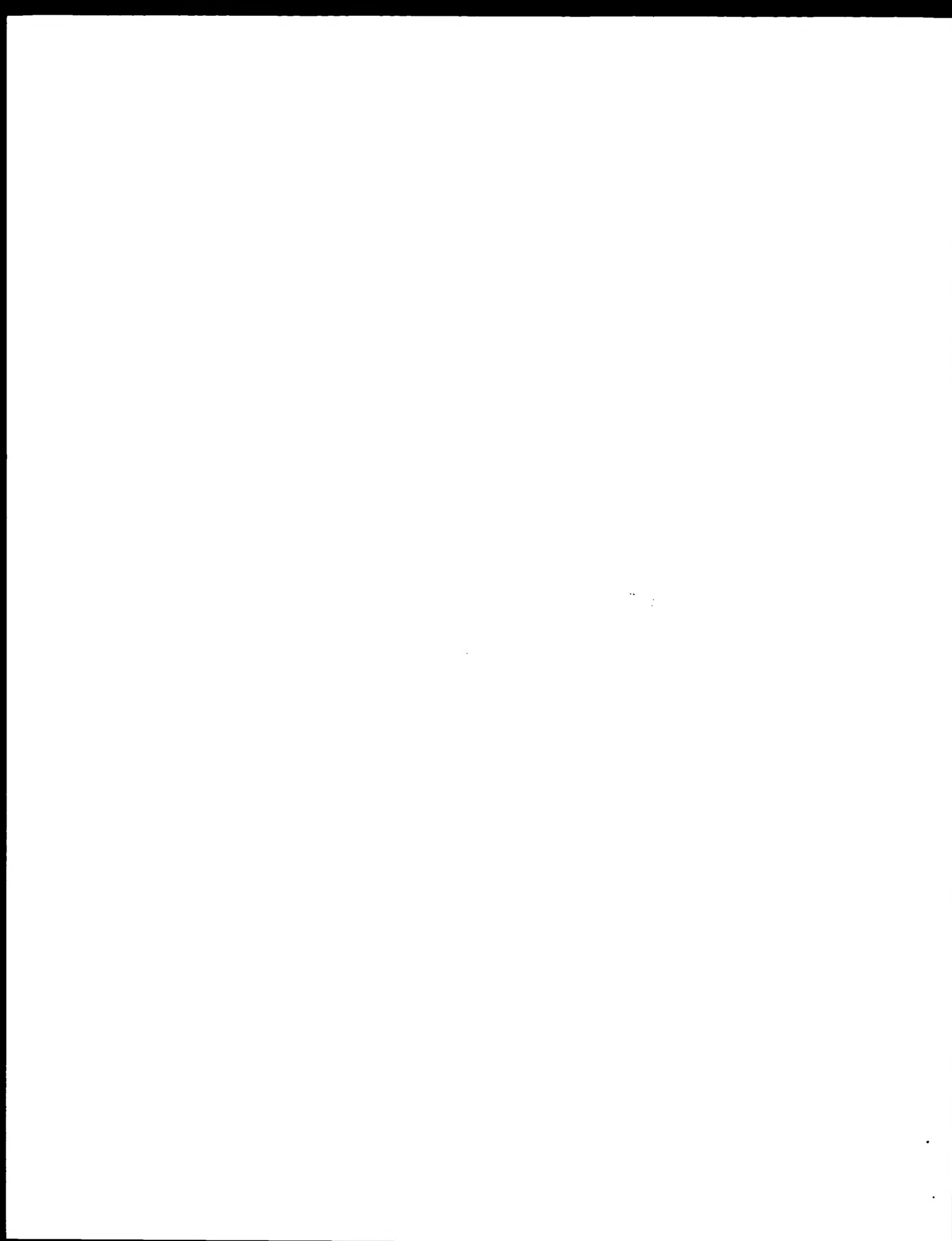
Query Match 11.0%; Score 198; DB 2; Length 4391.  
 Best Local Similarity 27.4%; Pred. No. 8-se 05;  
 Matches 93; Conservative 42; Mismatches 116; Indels 98; Gaps 20;

QY 5 TAPITGPEPLVKRGKAPKPPQPLEWKINTGRTFAWKVLSQAGSPWDSVAPVLPNGSLFL 54  
 Db 3032 IVUUGQVASHKCLTHGGAATSIEMKTKRNRLE-----DNV-HISPNGSI-I 3076  
 QY 65 PAVGIQ--DEGIFPCQAMPNGKEIKSNYPVAVYQIIPCKPREIVDSASLITAG-VPNKVG- 120  
 Db 3077 TIVGTPPSNIRTPVVASNAYG-VAGSVVNLNVH---GPPTV---SVIPEGPVWKVKGK 3128  
 QY 121 -----TCVSPGSGYPAGTILSW-HLDGKPLVPNEKGVSKYKTRRHPETGIIPTIOSMLMVTFA 175  
 Db 3129 AVTLECVSAGE-PRSSARKIRISSTP-----AKIIPRTY---GLMDSHAVIQISSA 3175  
 QY 176 RGGDPRPTFG-----SFSPCLPPIPALPTAPIQPRVWEVPVLEE 215  
 Db 3176 KPSP-ACTVYCLIAQNALGIAQKQVIVDTGAMACAVQ-----VQARE 3218  
 QY 216 VQLVWPEEGAVAPGGVILICDEVIAQVSIQIHKMKDGVVPIIP---PSVILIPPEIGIQ 272  
 Db 3214 AELTVE-----AHTATLP-SATISAPATIHWSKLSPLEWQHELESDTLIIPRAVQ 3271  
 QY 273 DQGTYSQVATHSHGHPQESRAVYSISITPEGEGPTAGSV 311  
 Db 3272 DSGQYICNAT-SPAGHAFA-----TIIHVESPPVATTV 3304

RESULT 15  
 150419  
 s-gicerin precursor - chicken  
 C:Species: Gallus gallus (chicken)  
 C>Date: 13-Sep-1996 #sequence\_revision 1<sequence\_change 21-101-2009  
 C:Accession: 150419  
 R:Taira, E.; Takahashi, N.; Taniura, H.; Kim, C.H.; Miki, N.  
 Neuron 12, 861-872, 1994  
 A>Title: Molecular cloning and functional expression of gicerin, a novel cell adhesion  
 A:Reference number: 150419; MIM:94213753; PMID:8161457  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-584 <TAI>  
 A:Cross-references: GR:038559; NID:g1009246; PID:HAA07563.1; PID:g559701

Query Match 11.0%; Score 194.5; DB 2; Length 584;  
 Best Local Similarity 26.8%; Pred. No. 1.5e-05;  
 Matches 66; Conservative 33; Mismatches 108; Indels 39; Gaps 10;

QY 52 SVARVL--PNGSIFDPAVGIDPHGTFPCQAMPNGKFTKSNYPVAVYQIIPCKPREIVDSAS 109



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2003, 15:37:16 ; Search time 19,4414 seconds  
(without alignments)  
708,288 Million cell updates/sec

Title: US-09-872-185B-2

Perfect score: 1766

Sequence: 1 AQNITARIPEVLKCKKAP .....GSHGATLALGILGLGTA 332

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SWISSProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1766	100.0	404	1	Q15109 homo sapien
2	1463	82.8	416	1	RAGE_HUMAN
3	1391	78.8	402	1	PAGE_BOVIN
4	1383	78.3	403	1	PAGE_PAT
5	1245	12.7	583	1	PAGE_MOUSE
6	219.5	12.4	583	1	C166_MOUSE
7	217.5	12.3	646	1	MU18_HUMAN
8	213	12.1	564	1	C166_BRARE
9	212.5	12.0	555	1	C166_CARAU
10	205.5	11.6	588	1	C166_CHICK
11	190.5	10.8	702	1	CEA5_HUMAN
12	189	10.7	439	1	PCBM_HUMAN
13	186.5	10.6	1914	1	KM15_HUMAN
14	185	10.5	515	1	PVRL_PIG
15	185	10.5	3707	1	PGBM_MOUSE
16	180.5	10.2	417	1	ICA5_FABIT
17	177	10.0	924	1	ICA5_HUMAN
18	175.5	9.9	444	1	NIP1_PAT
19	172.5	9.8	739	1	VZAL_PAT
20	172	9.7	1266	1	NACA_CHICK
21	170	9.6	1447	1	DPC_MOUSE
22	169.5	9.6	521	1	CEAL_MOUSE
23	168.5	9.5	515	1	PVRL_MOUSE
24	168	9.5	1051	1	PKR7_HUMAN
25	168	9.5	1070	1	PKR7_HUMAN
26	165	9.3	1040	1	AXO1_HUMAN
27	164	9.3	1447	1	DEC_HUMAN
28	163	9.2	517	1	PVRL_HUMAN
29	162	9.2	761	1	HCAC_HUMAN
30	162	9.2	848	1	NCA1_HUMAN
31	162	9.2	1234	1	NPHN_PAT
32	161	9.1	739	1	VZAL_HUMAN
33	160	9.1	837	1	NCA2_MOUSE

## RESULT 1

### RAGE\_HUMAN

AC Q15109; Q15279; Q9Y3R3; Q9H2X7; STANDARD; PRT: 404 AA.  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (rel. 41, Last annotation update)  
 DE Advanced glycosylation end product-specific receptor precursor  
 DE (Receptor for advanced glycosylation end products)  
 DE AGER OR RAGE.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RC SEQUENCE FROM N.A. (ISOFORM 1).  
 RP TISSUE=Lung;  
 RX MEDLINE=92340547; PubMed=1378843;  
 RA Neepser M., Schmidt A.M., Brett J., Yan S.D., Wang F., Pan Y.C.,  
 RA Elliston K., Stein D., Shaw A.;  
 RT Cloning and expression of a cell surface receptor for advanced  
 RT glycosylation end products of proteins.\*;  
 RL J. Biol. Chem. 267:14998-15004(1992).  
 RN [2]  
 RC SEQUENCE FROM N.A. (ISOFORM 1).  
 RP MEDLINE=95137587; PubMed=7835890;  
 RA Suqaya K., Fukagawa T., Matsumoto K., Mira K., Takahashi E., Ando A.,  
 RA Inoko H., Ikenura T.;  
 RT Three genes in the human MHC class III region near the junction with  
 RT the class II: gene for receptor of advanced glycosylation end  
 RT products, PBX2 homeobox gene and a notch homolog, human counterpart  
 RT of mouse mammary tumor gene int-3.\*;  
 RL Genomics 23:468-419(1994).  
 RN [3]  
 RC SEQUENCE FROM N.A. (ISOFORM 1).  
 RP Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,  
 RA Banta A., Spiles T., Hood L.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RC SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT ARG-100.  
 RA Abedin M.J., Yonekura H., Migita H., Karasawa J., Yamamoto Y.,  
 RA Yamamoto H.;  
 RT Molecular heterogeneity of the receptor for advanced glycation  
 RT endproducts.\*;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RC SEQUENCE FROM N.A. (ISOFORM 2).  
 RA Malherbe P., Richards J., Gaillard H., Thompson A., Diener C.,  
 RA Schuler A., Huber G.;  
 RT cDNA cloning of a novel secreted isoform of the human Receptor for  
 RT Advanced Glycation End products (RAGE) and characterization of cells  
 RT co-expressing cell-surface scavenger receptors and Swedish mutant  
 RT amyloid precursor protein.\*;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RC SEQUENCE FROM N.A. (ISOFORM 1).

P13594 mus musculus  
 P13595 mus musculus  
 P2063 rattus norv  
 P32004 homo sapien  
 Q9qzs7 mus musculus  
 P35331 gallus gall  
 O60500 homo sapien  
 Q90773 gallus gall  
 P28685 gallus gall  
 Q98919 gallus gall  
 P97798 mus musculus  
 Q60625 mus musculus

## ALIGNMENTS

TISSUE: Lung;  
RA Strassberg R.;  
KL Submitted (JAN 2002) to the EMBL/GenBank/DDBJ databases  
RN [7]  
RS  
RT SEQUENCE OF 1 12 FROM N.A.  
RA Hudson B.L., Futers T.S.,  
RT "Novel polymorphisms in the receptor for advanced glycation  
RT end products (RAGE) gene";  
RL Submitted (NOV 1999) to the EMBL/GenBank/DDBJ databases  
CC -! FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END  
CC PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS  
CC WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED  
CC RATE IN DIABETES  
CC -! SUBCELLULAR LOCATION: Type 1 membrane protein (isoform 1).  
CC Secreted (isoform 2).  
CC -! ALTERNATIVE PRODUCTS: 2 ISOMFORMS: 1 (SHOWN HERE) AND 2/RAGESEC;  
CC ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -! TISSUE SPECIFICITY: ENDOTHELIAL CELLS.  
CC -! SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY  
CC -! SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN LIKE V-TYPE DOMAIN.  
CC -! SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN LIKE C2-TYPE DOMAINS.  
CC  
CC this SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))  
CC  
CC EMBL: M91211; AAA03574.1;  
CC DR DR EMBL: D28769; BAA05958.1;  
CC DR DR EMBL: U89336; AAB47491.1;  
CC DR DR EMBL: AH046432; AAA89469.1;  
CC DR DR EMBL: A1138223; CAH43108.1;  
CC DR DR EMBL: BC020669; AAB20669.1;  
CC DR DR EMBL: AF208289; AAG34578.1;  
CC DR DR GenBank: U0001420; AGER.  
CC DR DR MIM: 600214;  
CC DR DR InterPro: IPR003006; Ig MH.  
CD DR DR InterPro: IPR003598; Ig C2.  
CD DR DR InterPro: IPR003600; Ig L1ke.  
CD DR DR Pfam: PF0047; Ig 2.  
CD DR DR SMART: SM00410; Ig L1ke; 1.  
CD DR DR SMART: SM00408; Ig C2; 1.  
CD DR DR PROSITE: PS00290; Ig MH; 1.  
CD DR DR Immunoglobulin domain; Glycoprotein; Repeat; Signal;  
CD KW Alternative splicing; Polymorphism.  
CD FT SIGNAL: 1 22  
CD FT CHAIN: 23 404  
CD  
CD DOMAIN 23 442  
CD FT TRANSMEM 443 463  
CD FT DOMAIN 464 404  
CD FT DOMAIN 41 106  
CD FT DOMAIN 147 215  
CD FT DOMAIN 252 308  
CD FT DISULFID 48 99  
CD FT DISULFID 144 208  
CD FT DISULFID 259 401  
CD FT CARBOHYD 25 25  
CD FT CARBOHYD 81 81  
CD FT DOMAIN 480 484  
CD FT VASPLITIC 54 67  
CD FT VASPLITIC 275 404  
CD  
CD VARIANI 100 100  
CD  
CD CONFLICT 1 1  
CD FT

SQ SEQUENCE 404 AA: 42802 MW: 6058404.460 kDa pI: 5.04  
 only Match: 100.0% Score: 1766; DB: 1; Length: 404  
 Best local Similarity: 100.0%; Pred. No. 8, 107  
 Matches 342; Conservative 1; Mismatches 0; Indels 0; Gaps 0

QY 1 AGNITAKGEPLVKKCGA-KKPPKURLEWKLNTGRTEAMKVLSPQGGHWSVAVKLLG 60  
 DB 23 AGNITAKGEPLVKKCGA-KKPPKURLEWKLNTGRTEAMKVLSPQGGHWSVAVKLLG 82  
 QY 61 SLFLPAVGIQERIGTRCQAMNPNCKPTKSNPVVYGIPTKRIEIVASSELTAGVNVK 120  
 DB 84 SLFLPAVGIQERIGTRCQAMNPNCKPTKSNPVVYGIPTKRIEIVASSELTAGVNVK 142  
 QY 121 TCVSGSPACTLSWHLLKGLVFNKGVSKVLTQRHHPTEGLTLOSLLMVTIPKGGP 180  
 DB 143 TCVSGSPACTLSWHLLKGLVFNKGVSKVLTQRHHPTEGLTLOSLLMVTIPKGGP 202  
 QY 181 RTTSCSTSGEGFRRHRLRTAFLEFVWEVPLEEVQLVVEEAWAVAKGLVLTSEVP 240  
 DB 204 RTTSCSTSGEGFRRHRLRTAFLEFVWEVPLEEVQLVVEEAWAVAKGLVLTSEVP 262  
 QY 241 AQPSPQIHWMKGVPLDLPSPVLITTEGHPQAGTYSVAHSSGNGQFSPAVSLSTE 300  
 DB 264 AQPSPQIHWMKGVPLDLPSPVLITTEGHPQAGTYSVAHSSGNGQFSPAVSLSTE 322  
 QY 301 PCEGPIAGSVGSGGLGIALALGILGLGIA 342  
 DB 324 PCEGPIAGSVGSGGLGIALALGILGLGIA 364

RESULT 2  
 RACE\_BOVIN  
 15 RACE\_BOVIN STANDARD; PRI: 416 AA.  
 AC Q28173;  
 DT 01 NOV 1997 (Rel. 45, Created)  
 DT 01 NOV 1997 (Rel. 45, Last sequence update)  
 DT 15 JUN 2002 (Rel. 41, Last annotation update)  
 DE Advanced glycosylation and product specific receptor precursor  
 DE (receptor for advanced glycosylation end products)  
 DE ACER FOR RACE.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Bovina; Bovidae;  
 OS Bovidae; Bovina; Bos.  
 OX NPRI:taxid:9913;  
 RN 111  
 RP  
 RC TISSUE: Lung;  
 RC MEDLINE: 92340547; PubMed:14740474;  
 RA Norpet M., Schmidt A.M., Brett J., Yan S.D., Wang Y., Fan Y.C.,  
 RA Elliston K., Stern D., Shaw A.J.  
 RA Cloning and expression of a cell surface receptor for advanced  
 RA glycosylation end products of proteins.  
 RA J. Biol. Chem. 267:14998-15014(1992).  
 CC 1 FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END  
 CC PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS  
 CC WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED  
 CC RATE IN DIABETES.  
 CC 1 SUBCELLULAR LOCATION: Type 1 membrane protein.  
 CC 1 TISSUE SPECIFICITY: ENDOTHELIAL CELLS.  
 CC 1 SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
 CC 1 SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN LIKE V TYPE DOMAIN  
 CC 1 SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN LIKE C2 TYPE DOMAINS

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/submit/submit.html>) or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)

```

DR EMBL: M91312; AAA03575.1; -.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig_2.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00408; IGG2; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal
FT SIGNAL 1 22
FT CHAIN 23 416
FT DOMAIN 23 352
FT TRANSMEM 353 373
FT DOMAIN 374 416
FT DOMAIN 31 105
FT DOMAIN 136 214
FT DOMAIN 262 318
FT DISULFID 38 98
FT DISULFID 143 207
FT DISULFID 269 311
FT CARBOHYD 25 25
FT CARBOHYD 80 80
FT DOMAIN 391 396
SQ SEQUENCE 416 AA; 44182 MW; 8703815573767AE CRC64;

Query Match 82.8%; Score 1463; DB 1; Length 416;
Best Local Similarity 81.5%; Pred No. 2.5e-87;
Matches 278; Conservative 21; Mismatches 30; Indels 12; Gaps 2;

QY 2 QNITARIQELPLVKCKGAPKPPQRLKNTGRTEAWKVLSPQGGPMDSVARVLPNGS 61
DB 24 QNITARIQELPLVKCKGAPKPPQRLKNTGRTEAWKVLSPQ-GDPWDSVARVLPNGS 82
QY 62 LFLPAVGQDEIFRCQAMNFKETKSNYRVVYQIPKPEIYDVSASELTAGVPNKVGT 121
DB 83 LLLPAVGQDEIFRCQAMNFKETKSNYRVVYQIPKPEIYDVSASELTAGVPNKVGT 142
QY 122 CVSEGSYFATLSWHLGDKLPLVPNEKGVSKVKEKTRHPETGLTQLQSELMVTTPARGGDP 181
DB 143 CVSEGSYPAGTLLNWLDDGKTLIPDGKGVSKVKEETKPKTKGLTGLHSELMTVPARGGALH 202
QY 182 PTFSCSPGLPRHRLRTAPIQPRVM-----EPVLEERVQIVVEPESGGAAPG 230
DB 203 PTFSCSPGLPRHRLRTAPIQPRVMSEHRRGSEPNVDVAVPLKVLVVEPESGGAAPG 262
QY 231 GTVTLTCEVAQSPQIHHMKDGVFLPLPSPVLLILPEIGSPQATYSVATHSSHGQPE 290
DB 263 GTVTLTCEVAQSPQIHHMKDGVFLPLPSPVLLILPEIGSPQATYSVATHSSHGQPE 322
QY 291 SRAVSTSTIBGEEPTASVWVSNGJSTLALALSLRLAT 331
DB 323 SRAVSVTIETGEGTTASVVEGPLETALTITAILGGLGT 363

RESULT 3
PAGE_PAT STANDARD: PRT; 402 AA.
AC Q63495;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, last sequence update)
DT 15-JUN-2002 (Rel. 41, last annotation update)
DE Advanced glycosylation end product-specific receptor precursor
DE (Receptor for advanced glycosylation end products).
GN AGER OR RAGE.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID-10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Lung;
RM MEDLINE-9736824; PUBMED-9221812.

```

Renard C, Chappey G, Wautier M.P., Nagashima M., Lundh E., Morser J., Zhao L., Schmidt A.M., Scherrenmann J.M., Wautier J.L.; "Recombinant advanced glycation end product receptor pharmacokinetics in normal and diabetic rats."; Mol. Pharmacol. 52:54-62(1997).

CC FUNCTION: MEDIATES INTERACTIONS OF ADVANCED GLYCOSYLATION END PRODUCTS (AGE). THESE ARE NONENZYMATICALLY GLYCOSYLATED PROTEINS WHICH ACCUMULATE IN VASCULAR TISSUE IN AGING AND AT AN ACCELERATED RATE IN DIABETES.

CC SUBCELLULAR LOCATION: Type I membrane protein.

CC TISSUE SPECIFICITY: ENDOTHELIAL CELLS.

CC SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.

CC SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.

CC SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L33413; AAA2027.1; -.
DR InterPro: IPR003006; Ig\_MHC.
DR InterPro: IPR003598; Ig\_C2.
DR Pfam: PF00047; Ig\_2.
DR SMART: SM00410; Ig\_Like; 1.
DR SMART: SM00408; IGG2; 1.
DR PROSITE: PS00290; IG\_MHC; 1.
KW Immunoglobulin domain; Glycoprotein; Transmembrane; Repeat; Signal.
FT SIGNAL 1 22
FT CHAIN 23 402
FT DOMAIN 23 341
FT TRANSMEM 342 362
FT DOMAIN 363 402
FT DOMAIN 31 105
FT DOMAIN 136 213
FT DOMAIN 250 306
FT DISULFID 38 98
FT DISULFID 143 206
FT DISULFID 257 299
FT CARBOHYD 25 25
FT CARBOHYD 80 80
FT SEQUENCE 402 AA; 42663 MW; 594481BC3A51E94E CRC64;

Query Match 78.8%; Score 1391; DB 1; Length 402;
Best Local Similarity 79.5%; Pred. No. 1e-82;
Matches 263; Conservative 24; Mismatches 42; Indels 2; Gaps 2.

QY 2 QNITARIQELPLVKCKGAPKPPQRLKNTGRTEAWKVLSPQGGPMDSVARVLPNGS 61
DB 24 QNITARIQELPLVKCKGAPKPPQRLKNTGRTEAWKVLSPQ-GDPWDSVARVLPNGS 82
QY 62 LFLPAVGQDEIFRCQAMNFKETKSNYRVVYQIPKPEIYDVSASELTAGVPNKVGT 121
DB 83 LLLPAVGQDEIFRCQAMNFKETKSNYRVVYQIPKPEIYDVSASELTAGVPNKVGT 142
QY 122 CVSEGSYFATLSWHLGDKLPLVPNEKGVSKVKEKTRHPETGLTQLQSELMVTTPARGGDP 181
DB 143 CVSEGSYPAGTLLNWLDDGKTLIPDGKGVSKVKEETKPKTKGLTGLHSELMTVPARGGALH 201
QY 182 PTFSCSPGLPRHRLRTAPIQPRVMSEHRRGSEPNVDVAVPLKVLVVEPESGGAAPG 241
DB 203 PTFSCSPGLPRHRLRTAPIQPRVMSEHRRGSEPNVDVAVPLKVLVVEPESGGAAPG 261
QY 242 QPSPIVHMKDGVFLPLPSPVLLILPEIGSPQATYSVATHSSHGQPE 301
DB 262 QPPQTHWIKDGTPLPLAPSPVLLPEVGHDEGIYSCVATHPSHSPQESPPVNRVET 321
QY 302 QEEGIAGSVGGSGLCILALALSLRLAT 332



CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC -----
DR EMBL; 095030; AAC06342.1; -.
DR EMBL; 125274; AAA37528.1; -.
DR HSSP; Q13740; 1KJC.
DR MGD; MGI:1313266; Alcam.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF06047; Ig; 3.
DR SMART; SM00409; Ig; 3.
DR PROSITE; PS00290; IG_MHC; FALSE_NEG.
KW Cell adhesion; Immunoglobulin domain; Glycoprotein; Transmembrane;
KW Repeat; Signal.
FT SIGNAL 1 27
FT CHAIN 28 583
FT DOMAIN 28 527
FT TRANSMEM 528 549
FT DOMAIN 528 583
FT DOMAIN 550 583
FT DOMAIN 56 120
FT DOMAIN 263 321
FT DOMAIN 347 399
FT DOMAIN 428 492
FT DISULFID 43 113
FT DISULFID 157 220
FT DISULFID 270 313
FT DISULFID 354 392
FT DISULFID 435 485
FT CARBOHYD 95 95
FT CARBOHYD 167 167
FT CARBOHYD 265 265
FT CARBOHYD 306 306
FT CARBOHYD 361 361
FT CARBOHYD 457 457
FT CARBOHYD 480 480
FT CARBOHYD 499 499
FT CONFLICT 427 454
FT CONFLICT 454 454
SQ SEQUENCE 583 AA; 65161 MW; E7BAPAF8FCA8B9489 CRC64;

Query Match
Best local similarity 25.98; Score 224.5; DB 1; Length 583,
Matches 79; Conservative 46; Mismatches 129; Indels 51; Gaps 12;

QY 4 ITARIGELVLKKGAPKPPQKLEWKLNTGRTEAKVLSPPQSSG-----PW 50
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
31 VNSAYGDTIVMPCK---LDVQNLMP---GK---WKYKPKDGPVFLAFRSSTKKSVOY 80
QY 51 DSVAP-----VLPNGSLFLPAVGIQDEGIFRCQAMNRNGKETKSNV-----RVRYVQIP 93
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
81 DDVPEYKDRLSLSENVYLSIANAKISDEKREVCMLV-----TEDNVFEAPTLVKYFKQP 134
QY 100 GKPIVDASFLTAGVINKVCTCVSGSYDAGTSLSHLDGKPLVPNEKGVSKVKEQTRRP 159
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
135 SKPEIVKAPFLETDLKLKGDICISRDSPDGNTWYRGVQLQPVGEVAILFKKEIDP 194
QY 160 ETGLFTLQSELMTVPARGSDPRPTFSCS---FSPGLPRHRLRTAPIQPRVWE-VVPLEE 215
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
195 CTQLYVTTSSILEYKFKTR-SDIQPFVTCVITYYGPS-----GQKTIYSEQIFDIYVPTQ 248
QY 216 VQWVEPEGGAAGGIVTLCGEVPAKPS-QLHMKKGAVPLPLPSPVLLILPEIGPQDQ 274
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
249 VTIVLPPKNAIKEGDNTILQCLNGNPPPEEFMYLPQGEIRSSNTYTLTDVVRNAT 308
QY 275 GTYSC 279
Db : : : :
309 GDYKC 313

PSSM 6
C166_HUMAN

```

```

ID AC Q13740; O60892; STANDARD; PRT; 583 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE C166 antigen precursor (Activated leukocyte-cell adhesion molecule)
DE (ALCAM).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=95279347; PubMed=7760007;
RA Bowen M.A., Patel D.B., Li X., Modrell B., Malacko A.R.,
RA Wang W.-C., Marquardt H., Neubauer M., Pesando J.M., Francke U.,
RA Haynes B.F., Aruffo A.;
RT "Cloning, mapping, and characterization of activated leukocyte-cell
adhesion molecule (ALCAM), a CD6 ligand.";
RL J. Exp. Med. 181:2213-2220(1995).
[2]
SEQUENCE OF 2-583 FROM N.A.
MEDLINE=98161327; PubMed=9502422;
RA Leyen W.G., van Kempen L.C., Gijzen E.G., van Groningen J.J.,
RA van Kooyk Y., Bloemers H.P., Swart G.W.;
RT "MEMD, a new cell adhesion molecule in metastasizing human melanoma
cell lines, is identical to ALCAM (activated leukocyte cell adhesion
molecule).";
RT Am. J. Pathol. 152:805-813(1998).
[3]
CD6-BINDING DOMAIN.
MEDLINE=96420463; PubMed=8823162;
RA Skonier J.E., Bowen M.A., Emswiler J., Aruffo A., Bajorath J.;
RT "Recognition of diverse proteins by members of the immunoglobulin
superfamily: delineation of the receptor binding site in the human
CD6 ligand ALCAM.";
RL Biochemistry 35:2287-12291(1996)
[4]
3D-STRUCTURE MODELING OF 28-133.
MEDLINE=96060095; PubMed=8520450;
RA Bajorath J., Bowen M.A., Aruffo A.;
RT "Molecular model of the N-terminal receptor-binding domain of the
human CD6 ligand ALCAM.";
RL Protein Sci. 4:1644-1647(1995).
CC -!- FUNCTION: CELL ADHESION MOLECULE THAT BINDS TO CD6. INVOLVED IN
NEURITE EXTENSION BY NEURONS VIA HETEROPHILIC AND HOMOPHILIC
INTERACTIONS. MAY PLAY A ROLE IN THE BINDING OF T AND B CELLS TO
ACTIVATED LEUKOCYTES, AS WELL AS IN INTERACTIONS BETWEEN CELLS OF
THE NERVOUS SYSTEM.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -!- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide C166 entry;
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd166.htm".
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
CC -----
DR EMBL; L38608; AAF59499.1; -.
DR EMBL; Y10183; CAA71256.1; -.
DR PDB; 1KJC; 03-APR-96.
DR GenBank; HGNC:400; ALCAM.
DR MIM; 601662;
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003600; Ig_MHC.
DR InterPro; IPR003600; Ig_Like.

```

```

DR 10000473 191 51
DR SMART: SM00409; IG: 2;
DR 1000410; IG: 130; 2;
DR PRE-SITE: PS00290; IG: 646; FALSE NEG.
KW Cell adhesion; Immunoglobulin domain; Glycoprotein; Transmembrane;
KW Repeat; Signal; 3D-structure; Polymorphism.
FT SIGNAL 1 27
FT CHAIN 28 583
FT DOMAIN 28 583 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 528 549 POTENTIAL.
FT DOMAIN 550 583 CYTOPLASMIC (POTENTIAL).
FT 36 120
FT DOMAIN 150 227 IG-LIKE V-TYPE DOMAIN 1.
FT DOMAIN 263 321 IG-LIKE C2-TYPE DOMAIN 2.
FT DOMAIN 347 399 IG-LIKE C2-TYPE DOMAIN 1.
FT DOMAIN 428 492 IG-LIKE C2-TYPE DOMAIN 3.
FT DISULFID 43 113 POTENTIAL.
FT DISULFID 157 220 POTENTIAL.
FT DISULFID 270 313 POTENTIAL.
FT DISULFID 354 392 POTENTIAL.
FT DISULFID 435 485 POTENTIAL.
FT CARBOHYD 91 91 N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 167 167 N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 480 480 N-LINKED (GLCNAC) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC) (POTENTIAL).
FT VARIANT 258 258 N-S.
FT VARIANT 401 401 /FT10-VAR 003907.
FT VARIANT M-S-T.
FT /FT10-VAR 003908.
SQ SEQUENCE 583 AA: 65142 MW: E024FB974A60284 CRC64:

Query Match 12.4%; Score 219.5; DB: 1; Length 583;
Best Local Similarity 25.9%; Pred. No. 30-07;
Matches 79; Conservative 45; Mismatches 130; Indels 51; Gaps 12;

QY 4 ITAKIGELVLKCKAPKPPGPIFWKINTPTFAWKVLSQGGV... ---PW 50
DB 31 VNSAVGCHITLDR...LVQGNMF...CK...WKYKPGDSIVFAPSSIKKSVY 80
QY 51 DSV...ARVLNGLSLFVAVGQDEITFCQAMNNGRETSNY...PVEVYQIP 99
DB 81 DAVPEYKORLNISENYTILSINARISDEKRFVCLV...TELVPEAPTIVKVEKOP 134
QY 100 GKPEIVDSASLEFACVPAKKVGLVSEGGYVACITLSSWHLGGADVGNKAGSVAKGIRHP 159
DB 145 SKPEIVSVAIFFTEPLFKEI...ISEDSYDNTWYNGVLIHLEAVVITFKKEMDF 194
QY 160 EIGLILSELMLVIFARAGDEPTSS...ESPQIPPRALPTAPTPPVEV...PVELEE 215
DB 195 VTLQVMTSLTPEYKTRADLMPDECSVTVYGD...CQKTHSCQAVGVLYITPO 248
QY 216 VQLVVEPEGAVAPGSHVITVEFAQSPQIHWKIAVPLPLPSPVLLPEIGPQDQ 274
DB 249 VTIQVLPKRNALKEGDNITILKLGNGNPPPEETFLYFGQPEGIRSSNTYTLMDVRNAT 308
QY 275 GPYSC 279
DB 309 GLYK 314

RESULT 7
M18 HUMAN
AC E4321;
DT 01 NOV 1995 (rel. 32, created)
DT 01 NOV 1995 (rel. 42, last sequence update)
DT 15 JUN 2002 (rel. 41, last annotation update)
DE Cell surface glycoprotein Muc18 precursor (melanoma-associated antigen

```

```

DE MUC18 (Melanoma associated antigen A32) (35 endothelial
DE associated antigen) (CD146 antigen) (Melanoma adhesion molecule).
GN MCAM OR MUC18.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NX NCHI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-Melanoma;
RX MEDLINE=9009368; PubMed=2602481;
RA Lehmann J.M., Riethmuller G., Johnson J.P.;
RT "MUC18, a marker of tumor progression in human melanoma, shows
RT sequence similarity to the neural cell adhesion molecules of the
RT immunoglobulin superfamily.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:9891-9895(1989).
RN [2]
RP REVISIONS, SEQUENCE FROM N.A.
RX TISSUE-Melanoma;
RX MEDLINE=93391384; PubMed=8378324;
RA Sers C., Kirsch K., Rothbacher U., Riethmuller G., Johnson J.P.;
RT "Genomic organization of the melanoma-associated glycoprotein MUC18:
RT implications for the evolution of the immunoglobulin domains.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:8514-8518(1993).
RN [3]
RP SEQUENCE OF 24-44; 98-112; 136-166; 240-260; 379-389 AND 460-478.
RX MEDLINE=94215196; PubMed=8162602;
RA Shih L.M., Elder D.E., Speicher D., Johnson J.P., Herlyn M.;
RT "Isolation and functional characterization of the A32 melanoma
RT associated antigen.";
RL Cancer Res. 54:2514-2520(1994).
RN [4]
RP SEQUENCE OF 27-49; 98-112 AND 246-260.
RX MEDLINE=96136302; PubMed=8574133;
RA Bardio N., Frances V., Lesauve G., Ibrschowski N., George F.,
RA Sempol J.;
RT "Identification of the S Endo 1 endothelial associated antigen";
RL Biochem. Biophys. Res. Commun. 218:210-216(1996).
RN [5]
RP FUNCTION.
RX MEDLINE=9412524; PubMed=829890;
RA Johnson J.P., Rothbacher U., Sers C.;
RT "The progression associated antigen MUC18: a unique member of the
RT immunoglobulin supergene family.";
RL Melanoma Res. 3:337-340(1993).
CC -1- FUNCTION: COULD BE AN ADHESION MOLECULE ACTIVE IN NEURAL CREST
CC CELLS DURING EMBRYONIC DEVELOPMENT. ITS EXPRESSION MAY ALLOW
CC MELANOMA CELLS TO INTERACT WITH CELLULAR ELEMENTS OF THE VASCULAR
CC SYSTEM THEREBY ENHANCING HEMATOGENOUS TUMOR SPREAD.
CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.
CC -1- TISSUE SPECIFICITY: MAY APPEAR AT THE SURFACE OF NEURAL CREST
CC CELLS DURING THEIR EMBRYONIC MIGRATION. APPEARS TO BE LIMITED TO
CC VASCULAR SMOOTH MUSCLE IN NORMAL ADULT TISSUES. ASSOCIATED WITH
CC TUMOR PROGRESSION AND THE DEVELOPMENT OF METASTASIS IN HUMAN
CC MALIGNANT MELANOMA. EXPRESSED MOST STRONGLY ON MELANOCYTIC LESIONS
CC AND ADVANCED PRIMARY TUMORS AND IS ONLY RARELY DETECTED IN BENIGN
CC MELANOCYTIC NEVI. BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN LIKE C2 TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN LIKE V TYPE DOMAINS.
CC -1- DATABASE: NAME=PROV; NCHI=CD guide CD146 entry
CC WWW: http://www.ncbi.nlm.nih.gov/Entrez/Entrez.fcgi?db=pubmed
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement! (See http://www.ebi.ac.uk/submit/
CC or send an email to license@ebi.ac.uk)
CC EMBL: M29277; AAA:0824.1;

```



EMBL: M28882; AAA20922.1; ...  
 DR EMBL: X68254; CAA48332.1; JOINED.  
 DR EMBL: X68265; CAA48332.1; JOINED.  
 DR EMBL: X68266; CAA48332.1; JOINED.  
 DR EMBL: X68267; CAA48332.1; JOINED.  
 DR EMBL: X68268; CAA48332.1; JOINED.  
 DR EMBL: X68270; CAA48332.1; JOINED.  
 DR EMBL: X68271; CAA48332.1; JOINED.  
 DR Genew: HGNC:6934; MCAM.  
 DR MIN: 155735; ...  
 DR InterPro: IPR000406; Ig\_MHC.  
 DR InterPro: IPR003598; Ig\_C2.  
 DR InterPro: IPR003600; Ig\_Like.  
 DR Pfam: PF00047; Ig\_5.  
 DR SMART: SM00410; Ig\_Like\_2.  
 DR SMART: SM00408; Ig\_2; 2  
 KW Cell adhesion; Immunoglobulin domain; Glycoprotein; Transmembrane;  
 KW Repeat; Signal;  
 FT SIGNAL: 1 23  
 FT CHAIN: 24 646 CHIL SURFACE GLYCOPROTEIN MUC1A  
 FT DOMAIN: 24 559 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 560 583 POTENTIAL.  
 FT DOMAIN: 584 646 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN: 41 123 IG-LIKE V-TYPE DOMAIN 1.  
 FT DOMAIN: 154 230 IG-LIKE V-TYPE DOMAIN 2.  
 FT DOMAIN: 255 327 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN: 358 414 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DOMAIN: 445 506 IG-LIKE C2-TYPE DOMAIN 3.  
 FT DISULFID 48 116 PROBABLE.  
 FT DISULFID 161 223 PROBABLE.  
 FT DISULFID 272 329 PROBABLE.  
 FT DISULFID 365 407 PROBABLE.  
 FT DISULFID 452 499 PROBABLE.  
 FT CARBOHYD 56 56 N-LINKED (GLNAC...) (POTENTIAL).  
 FT CARBOHYD 418 418 N-LINKED (GLNAC...) (POTENTIAL).  
 FT CARBOHYD 449 449 N-LINKED (GLNAC...) (POTENTIAL).  
 FT CARBOHYD 467 467 N-LINKED (GLNAC...) (POTENTIAL).  
 FT CARBOHYD 508 508 N-LINKED (GLNAC...) (POTENTIAL).  
 FT CARBOHYD 518 518 N-LINKED (GLNAC...) (POTENTIAL).  
 FT CARBOHYD 527 527 N-LINKED (GLNAC...) (POTENTIAL).  
 FT CARBOHYD 544 544 N-LINKED (GLNAC...) (POTENTIAL).  
 SQ SEQUENCE 646 AA, 71793 MW, F064A5DAER0BARC6 CRG64.  
 Query Match 12.3%; Score 217.5; DR 1; Length 646.  
 Best Local Similarity 23.3%; Pred. No. 4.5e-07;  
 Matches 95; Conservative 55; Mismatches 135; Indels 123; Gaps 19;  
 QY 4 ITATIGPLVLKCKGAHKPP-QLRW--KLNTGRTEAMKVLSPQG---GGPWDSVARVL 57  
 DB 36 VEVHVGSTALIKGLSUSQNLSHVIMF-SVHAKKRIIFPVVQDQDSKSPGSEYKRLSLD 95  
 QY 58 PNR-SLELPVAVTIDRITFQVAMNIN-KETKS-NYPV--PVYQTPSKPEI-----V 105  
 DB 96 DRGALLALQVLPQDRIKFIQV-----GRPKRSQYRQIYRYKAPENIQVNIPLGIPV 150  
 QY 106 DSASELTAGVNVKAGTGVSGSYDPATISWHLISKPIVINKEGVSVKPKIIPHPHPLT 165  
 DB 151 NSKE-----PEVATCVGPNYPIQVVIWYNGKRLKEEKNRVIH-USSQTVSSGLYT 203  
 QY 166 IQSHLVTPAGQDPPTFSCFSPGLPRHPLR-----TADP---QPRVW ETVPL- 213  
 DB 204 IQSLIKALVK-EKUAQFYCLNYRPLPSGNHMKESREVTVPVFPYPTKVLVEVPVGM 262  
 QY 214 -----EE----- 215  
 DB 263 KEGDPVPIPCIANVNDPHPSISKQNPSTREAREETTINGVILVLEPAKHSRGYEQCA 322  
 QY 216 -----VQLVVPF-----CAVAP-----GVVILRQVPAQSPQTHMKIKV 254  
 DB 323 WLLDTMLISLSEFQELLVNVSVKVSFAAPFQVRSISLTLTCESSNQLEFQVLRREET 362  
 QY 255 PLPLPPSPVLLDIEIGHQDGIYSQVAT-HSSHGPFQFSKAVSISTEP 301

100 383 LGVIFKSPVLEQLHHLKPPAAGGYSYKCVASVPSIPGIPNIGLVKLAIFGP 440  
 RESULT 8  
 C166\_RRAPE  
 ID C166\_RRAPE STANDARD; PRT: 564 AA.  
 AC Q90460; Q90480;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE C166 antigen homolog precursor (Neurolin) (NM-GRASP homolog).  
 GN C166.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID:7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 FX MEDLINE:94376084; PubMed:80896660;  
 RA Karki J.P., Chang S., Kuwada J.Y.;  
 RT "The molecular cloning and characterization of potential chick  
 DM-GRASP homologs in zebrafish and mouse.";  
 RL J. Neurobiol. 25:831-845(1994).  
 RN [2]  
 RP SEQUENCE OF 398-561 FROM N.A.  
 RX MEDLINE:94299340; PubMed:8026643;  
 RA Laessing U., Giordano S., Strecher R., Lettspeich F., Stuermer C.A.;  
 RT "Molecular characterization of fish neurolin: a growth-associated  
 RT cell surface protein and member of the immunoglobulin superfamily in  
 the fish retinotectal system with similarities to chick protein  
 DM-GRASP/SC-1/HER.";  
 RL Differentiation 56:21-29(1994).  
 CC -|- FUNCTION: CELL ADHESION MOLECULE. INVOLVED IN NEURITE EXTENSION BY  
 CC NEURONS VIA HETEROPHILIC AND HOMOPHILIC INTERACTIONS.  
 CC -|- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -|- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
 CC -|- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC -|- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 DR EMBL: L25273; AAA50024.1; ...  
 DR EMBL: L25057; AAA50048.1; ...  
 DR HSP: Q13740; IG\_C2;  
 DR ZFIN: ZDB-GENE-990415-40; ccl66  
 DR InterPro: IPR003006; Ig\_MHC.  
 DR InterPro: IPR003600; Ig\_Like.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig\_4.  
 DR SMART: SM00410; Ig\_Like; 3.  
 DR SMART: SM00406; Ig\_V; 1.  
 DR PROSITE: PS00240; Ig\_MHC; 1  
 KW Cell adhesion; Immunoglobulin domain; Glycoprotein; Transmembrane;  
 KW Repeat; Signal;  
 FT SIGNAL: 1 24 POTENTIAL.  
 FT CHAIN: 25 564 C166 ANTIGEN HOMOLOG.  
 FT DOMAIN: 25 507 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 508 528 POTENTIAL.  
 FT DOMAIN: 529 564 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN: 31 117 IG-LIKE V-TYPE DOMAIN 1.  
 FT DOMAIN: 147 232 IG-LIKE V-TYPE DOMAIN 2.  
 FT DOMAIN: 256 314 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN: 340 389 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DOMAIN: 419 477 IG-LIKE C2-TYPE DOMAIN 3.  
 FT DISULFID 38 110 POTENTIAL.  
 FT DISULFID 154 217 POTENTIAL.



ID C166-CHICK STANDARD; PRI: 586 AA.  
 AC P42263;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE C166 antigen precursor (Swi glycoprotein) (BEN glycoprotein) (PM-  
 DE GWASP protein) (JG7 protein).  
 OS Gallus gallus (Chicken).  
 OS Gallus gallus (Chicken).  
 OC Archaeopteryx, Metazoa, Chordata, Craniata, Vertebrata, Palaeostomi;  
 OC Archosauria, Aves, Neognathae, Galliformes, Phasianidae, Phasianinae;  
 OC Gallus.  
 UN NCBI\_Linxlin-6041;  
 UN [1]  
 PN SEQUENCE FROM N A , AND SEQUENCE OF 34-53  
 PP TISSUE=Embryo;  
 RX MEDLINE=92010150; PubMed 191849;  
 RX Tanaka H., Matsui I., Adachi K., Tomura M., Kubota I.,  
 RX McFarland K.C., Kohr B., Lee A., Phillips H.S., Shelton D.L.;  
 RA "Molecular cloning and expression of a novel adhesion molecule, Scl-1,"  
 RA Neuron 7:545-545(1991)  
 RL [2]  
 RN SEQUENCE FROM N A , AND SEQUENCE OF 14-52, 87-94, 104-112, AND 504-502  
 RP TISSUE=Bursa of Fabricius;  
 RC MEDLINE=9240224; PubMed 160903;  
 RA Pourquie O., Corbel C., Le Gall J.-P., Kossler J., Le Douarin N.M.;  
 RA "HEN, a surface glycoprotein of the immunoglobulin superfamily, is  
 RA expressed in a variety of developing systems,"  
 RL Proc. Natl. Acad. Sci. U S A 89:5261-5265(1992)  
 RN POSSIBLE FUNCTION.  
 PP MEDLINE=9221111; PubMed 1313197;  
 RA Pourquie O., Hannonet M.E.R., Le Douarin N.M.;  
 RA "Association of BEN glycoprotein expression with climbing fiber  
 RA axogenesis in the avian cerebellum,"  
 RL J. Neurosci. 13:1548-1557(1993)  
 CC -1- FUNCTION: HOMOPHILIC ADHESION MOLECULE ASSOCIATED WITH CLIMBING  
 CC FIBER AXONogenesis. SUPPORTS NEURITE EXTENSION.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A RESTRICTED POPULATION OF AXONS.  
 CC WITHIN THE SPINAL CORD IT IS LOCALIZED TO AXONS IN THE DORSAL  
 CC FUNICULUS, MIDLINE FLOOR PLATE CELLS, AND MOTO-NEURONS. FOUND IN  
 CC EPITHELIA, CENTRAL NERVOUS SYSTEM, AND HEMOPOIETIC CELLS. EACH  
 CC DEVELOPING SYSTEM CARRIES DIFFERENTLY GLYCOSYLATED MOLECULAR FORMS  
 CC OF BEN.  
 CC -1- DEVELOPMENTAL STAGE: WIDELY EXPRESSED DURING EMBRYONIC  
 CC DEVELOPMENT.  
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.  
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE V-TYPE DOMAINS.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation at  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMRL: S63276; AAB20170.1;  
 EMBL: M76678; AAA48602.1;  
 EMRL: X64301; CAA45579.1;  
 DR HSP; Q13740; 1KJG.  
 DR InterPro: IPR003599; Iq.  
 DR InterPro: IPR003006; Iq\_MHC  
 DR InterPro: IPR003600; Iq\_Like

DR Plan: PR00047; Iq; 5;  
 DR SMART: SM00409; Iq; 3;  
 DR SMART: SM00410; Iq\_Like; 2;  
 DR PROSITE: PS00290; Iq\_MHC; FALISE\_NHG.  
 KW Cell adhesion. Immunoglobulin domain, Glycoprotein, Transmembrane,  
 KW Repeat: Signa...  
 FT SIGNAL 1 33  
 FT CHAIN 34 588  
 FT DOMAIN 34 532  
 FT TRANSMEM 533 553  
 FT DOMAIN 554 588  
 FT DOMAIN 42 126  
 FT DOMAIN 136 233  
 FT DOMAIN 259 326  
 FT DOMAIN 352 404  
 FT DOMAIN 433 497  
 FT DISULFID 49 119  
 FT DISULFID 163 226  
 FT DISULFID 276 319  
 FT DISULFID 359 397  
 FT DISULFID 440 490  
 FT CARBOHYD 101 101  
 FT CARBOHYD 173 173  
 FT CARBOHYD 194 194  
 FT CARBOHYD 271 271  
 FT CARBOHYD 312 312  
 FT CARBOHYD 366 366  
 FT CARBOHYD 462 462  
 FT CARBOHYD 485 485  
 FT CARBOHYD 504 504  
 FT CONFLICT 1 10  
 FT CONFLICT 25 25  
 FT CONFLICT 112 113  
 FT CONFLICT 329 329  
 FT CONFLICT 401 402  
 FT CONFLICT 401 402  
 SQ SEQUENCE 586 AA, 6572c MW, 2A28612D0164531E CRC64,  
 Query Match 11.6%, Score 205.5, DB 1, Length 588;  
 Best local Similarity 24.8%, Pred. No. 2.4e-06;  
 Matches 75; Conservative 46; Mismatches 135; Indels 47; Gaps 11;  
 QY 4 ITARIGELIVLKCKGAPKPPQRLWK--LMTGTEAKVKVLSPOGG----- 48  
 DB 37 VNAVYGDITIMPC-----PDEVPGMLPGR--WKYEMPSSPVFTAFSSSTKKNV 84  
 QY 49 PWNSV-----APVLPNSLFLPAV5IGUEGIFEDAMNPNKCKTKSVFVYVQIPKRP 102  
 LB 65 QYDQVDEYKDRPESLSENVTLSIKNAPISDEKPFVMLVTEDDVSEPTV-VKVFQKPSQP 143  
 QY 103 EIVDSASELTAGVPKNGVGVSEGSYPASTISWHLGKPLVPNEKGVSKVKEGTPPHPETG 162  
 DB 144 EILHCADFELETEKLMGLGECVVRSYPGNVTWYKNGRVLQPEVVVNLKRVENRSTG 203  
 QY 163 LETLSELMVTIPARGDPPPTFSQSPQLPPHALPTAPICPPVWE-PVPLEEVQLVWE 221  
 LB 204 LEIMTSSLYQYMEIK-EDANAKETCLIVYHSGG--KTIQSEPVYFDVHYPTKVTIRVL 260  
 QY 222 PRGGVAVPGGVITVTCFVPAQPSPO-----IHWKMDGVPLPLPPLPILPFIQPOOGT 276  
 LB 261 SSSSIKEGDNVTLKLSGNGNPPPPUEFLFYIPGETEGL----RSSDIYVMTDVRNATGE 316  
 QY 277 YSC 279  
 DB 317 YKC 319  
 RESULT 11  
 CEAS\_HUMAN  
 ID CEAS\_HUMAN STANDARD; PRI: 702 AA.  
 AC P06731;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)



QY 111 -----AGVPKNVGTGCVS 124  
 DB 345 CEPEIQNTYLMWVNNGLSPVSPRLQSLNDNRRLTLSTVRNDVCPYEGIGQNELSVDSHS 404  
 QY 125 E-----GSPYAGTSLWILLUCKPLVPNEKGV 150  
 DB 405 DPVILNVLGDDPTLSPSYVYRPGVNLISLSCHAAASNPAAQYSLWIDG-----N 454  
 QY 151 VKEGTR-----HPEITGLILQSELMTVPARGD-----PRFTFSCSPSP 190  
 DB 455 IQQHTQELFTSNIETKNSGLYTQOAN---NSASGHSRTTIVTKITVSABLKPESISSNNSK 511  
 QY 191 GLPHRAL-----RTAPIQPRVW----- 208  
 DB 512 PVEDKDAVAFTEPEAGNTYLMWVNGSLPVSPLQLISNGNRITLTLENTVNDARAYVC 571  
 QY 209 -----EVPLEEV-----QLVVEPEGAVAPGGTTLTCEVPAQSPQIIHWKMDG 253  
 DB 572 GIONSVANSKSDPVLIDVLYGDPPIISPDSYLSGANINLSCHASNPSPQYSWRING 631  
 QY 254 VPLPSPVLLIFEPDQATYSCVATHSHGHPQESRAVSIETIEPCEGPTAGSVGG 313  
 DB 632 I--PQHTQVLFIAKTIPNNNTYA--FVSNLATGRNNSIVKSIIVSASGT-----SPGL 683  
 QY 314 SGHCTLALALGILGGL 329  
 DB 684 SAGATVGMIGLVGV 699

## RESULT 12

PGBM\_HUMAN  
 ID PGBM\_HUMAN STANDARD; PRT: 4393 AA.  
 AC p98160; Q16287;  
 DT 01-OCT-1996 (Rel. 34, created)  
 DT 01-OCT-1996 (Rel. 34, last sequence update)  
 DT 15-JUN-2002 (Rel. 41, last annotation update)  
 DE Basal membrane specific heparan sulfate proteoglycan core  
 DE protein precursor (HSPG) (Perlecan) (PLC).  
 GN HSPG2.  
 OS Homo sapiens (Human).  
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;  
 OC Mammalia, Eutheria, Primates, Catartini, Hominoidea, Homo.  
 CX NCBI TaxID:9606;  
 RN [1]  
 SEQUENCE FROM N.A.  
 EX MEDLINE-92112994; PubMed-1730768.  
 RA Kallunki P., Tytgavason K.;  
 RT "Human basement membrane heparan sulfate proteoglycan core protein: a  
 RT 467-kD protein containing multiple domains resembling elements of the  
 RT low density lipoprotein receptor, laminin, neural cell adhesion  
 RT molecules, and epidermal growth factor.";  
 RL J. Cell Biol. 116:559-571(1992).  
 RN [2]  
 SEQUENCE FROM N.A.  
 RC TISSUE-Skin, and Colon;  
 RX MEDLINE-92246394; PubMed-1549182;  
 RA Murdoch A.P., Dodge G.P., Cohen I., Tuan R.S., Iozzo R.V.;  
 RT "Primary structure of the human heparan sulfate proteoglycan from  
 RT basement membrane (HSPG2/perlecan): A chimeric molecule with multiple  
 RT domains homologous to the low density lipoprotein receptor, laminin,  
 RT neural cell adhesion molecules, and epidermal growth factor.";  
 RL J. Biol. Chem. 267:8544-8557(1992).  
 RN [3]  
 SEQUENCE OF 1018-1472 FROM N.A.  
 RC TISSUE-Colon;  
 RX MEDLINE-91365376; PubMed-1679749;  
 RA Dodge G.R., Knäuper J., Chi M.L., Hassell J.P., McRide G.W.,  
 RA Yi H.F., Iozzo R.V.;  
 RT "Heparan sulfate proteoglycan of human colon: partial molecular  
 RT cloning, cellular expression, and mapping of the gene (HSPG2) to the  
 RT short arm of human chromosome 1.";  
 RL Genomics 10:673-680(1991).  
 RN [4]

SEQUENCE OF 892-1398 FROM N.A.  
 RC TISSUE-Fibroblasts;  
 RX MEDLINE-92110660; PubMed-1685141;  
 RA Kallunki P., Eddy R.L., Byers M.G., Kestila M., Shows T.B.,  
 RA Tytgavason K.;  
 RT "Cloning of human heparan sulfate proteoglycan core protein,  
 RT assignment of the gene (HSPG2) to 1p36.1--p35 and identification of  
 RT a BamHI restriction fragment length polymorphism";  
 RL Genomics 11:389-396(1991).  
 RN [5]  
 SEQUENCE OF 1-21 FROM N.A.  
 RX MEDLINE-94052171; PubMed-8234307;  
 RA Cohen I.R., Graessle S., Murdoch A.D., Iozzo R.V.;  
 RA "Structural characterization of the complete human perlecan gene and  
 RA its promoter.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:10404-10408(1993).  
 CC -1- FUNCTION: This protein is an integral component of basement  
 CC membranes. It is responsible for the fixed negative electrostatic  
 CC charge and is involved in the charge-selective ultrafiltration  
 CC properties. It serves as an attachment substrate for cells.  
 CC -1- SUBUNIT: Purified perlecan has a strong tendency to aggregate in  
 CC dimers or stellate structures. It interacts with other basement  
 CC membrane components such as laminin, prolargin and collagen type  
 CC IV.  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES.  
 CC -1- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED  
 CC AND O-LINKED OLIGOSACCHARIDES.  
 CC -1- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 10.5 LAMININ EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 LAMININ DOMAINS IV.  
 CC -1- SIMILARITY: CONTAINS 22 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 4 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 SEA DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X62515; CA444373.1;  
 DR EMBL: M85289; AA52700.1;  
 DR EMBL: M54283; AA52699.1;  
 DR EMBL: S75436; AA831121.2;  
 DR EMBL: L22078; NOT\_ANNOTATED\_CDS.  
 DR HSPG: P00740; 1EDM.  
 DR Sleno-2DPAGE; P9E160;  
 DR Genew; HGNC:5273; HSPG2.  
 DR MIM; 142461;  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF\_2.  
 DR InterPro: IPR001438; EGF\_11.  
 DR InterPro: IPR003406; Iq\_MHC.  
 DR InterPro: IPR003598; Iq\_C2.  
 DR InterPro: IPR002172; LDL\_recept\_A.  
 DR InterPro: IPR000634; Laminin\_B.  
 DR InterPro: IPR002849; Laminin\_EGF.  
 DR InterPro: IPR001791; Laminin\_G.  
 DR InterPro: IPR000882; SEA\_domain.  
 DR Pfam: PF00008; EGF; 4.  
 DR Pfam: PF00047; Iq; 22.  
 DR Pfam: PF00052; laminin\_B; 3.  
 DR Pfam: PF00053; laminin\_EGF; 7.  
 DR Pfam: PF00054; laminin\_G; 3.  
 DR Pfam: PF00057; ldl\_recept\_a; 4.  
 DR Pfam: PF01390; SEA; 1.  
 DR PRINTS: PR00010; EGFBL00D.  
 DR ProDom: PD003C31; Laminin\_B; 3.  
 DR SMART; SM0018C; EGF\_Lam; 6.



RC TISSUE-Umbilical vein;  
 RX MEDLINE-99216419; PubMed-10198165;  
 PA Lazar V.I., Garcia J.G.N.;  
 RT "A single human myosin light chain kinase gene (MLCK; MYLK).";  
 PL Genomics 57:266-267(1996)  
 [4]  
 RN REVISIONS (ISOFORM 2);  
 RA Hirukov K.G., Garcia J.G.N.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [5]  
 RN SEQUENCE OF 923-1914 FROM N.A.  
 RP TISSUE-Hippocampus;  
 RX MEDLINE-96121365; PubMed-8575746;  
 RA Potter M.C., Cholet E., Pekarsky Y., Gardiner K., Rossier J.,  
 PA Turnell W.G.;  
 RT "The human myosin light chain kinase (MLCK) from hippocampus:  
 RT cloning, sequencing, expression, and localization to 3qter-q21";  
 RL Genomics 29:562-570(1995).  
 [6]  
 RN SEQUENCE OF 1614-1914 FROM N.A.  
 RP TISSUE-Lung, and Placenta;  
 RX MEDLINE-20007938; PubMed-10536370;  
 PA Waterson M., Schavoy J.P., Guo L., Weiss G., Chlenski A.,  
 RA Shinsky V.P., Van Eldik L.J., Harech J.;  
 RT 3q21 in a multi gene cluster: organization, expression, alternative  
 RT splicing and polymorphic marker";  
 RL J. Cell Biochem 75:481-491(1999).  
 [7]  
 RN SEQUENCE OF 1456-1914 FROM N.A.  
 RP TISSUE-Placenta;  
 RA Waterson M.D.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases  
 [8]  
 CC SMOOTH MUSCLE CONTRACTION VIA PHOSPHORYLATION OF A SPECIFIC SERINE  
 CC IN THE N-TERMINUS OF MYOSIN LIGHT CHAINS (MLC), AN EVENT THAT  
 CC FACILITATES MYOSIN INTERACTION WITH ACTIN FILAMENTS. CENTRAL  
 CC DETERMINANT IN THE DEVELOPMENT OF VASCULAR PERMEABILITY AND TISSUE  
 CC EDEMA FORMATION. IN THE NERVOUS SYSTEM IT HAS BEEN SHOWN TO  
 CC CONTRIBUTE TO THE GROWTH INITIATION OF ASTROCYTES IN CULTURE  
 CC AND TO PARTICIPATE IN TRANSMITTER RELEASE AT SYNAPSES FORMED  
 CC BETWEEN CULTURED SYMPATHETIC GANGLION CELLS. CRITICAL PARTICIPANT  
 CC IN SIGNALING SEQUENCES THAT RESULT IN FIBROBLAST APOPTOSIS.  
 CC [9]  
 CC CATALYTIC ACTIVITY: ATP + [myosin light-chain] = ADP + [myosin  
 CC light-chain] phosphate.  
 CC [10]  
 CC SUBUNIT: TELOKIN BINDS CALMODULIN.  
 CC [11]  
 CC ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS: A NON-MUSCLE FORM (THE  
 CC LONGEST FORM), A SMOOTH-MUSCLE FORM AND TELOKIN (A C-TERMINAL  
 CC SECTION WITH NO CATALYTIC ACTIVITY); ARE PRODUCED BY ALTERNATIVE  
 CC INITIATION.  
 CC [12]  
 CC ALTERNATIVE PRODUCTS: AT LEAST 6 ISOFORMS: 1 (SHOWN HERE), 2, 3A,  
 CC 3B, 4 AND DEL-1790; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC [13]  
 CC TISSUE SPECIFICITY: SMOOTH MUSCLE AND NON-MUSCLE ISOZYME ARE  
 CC EXPRESSED IN A WIDE VARIETY OF ADULT AND FETAL TISSUES AND IN  
 CC CULTURED ENDOTHELIUM WITH QUALITATIVE EXPRESSION APPEARING TO BE  
 CC NEITHER TISSUE- NOR DEVELOPMENT-SPECIFIC. NON-MUSCLE ISOFORM 2 IS  
 CC THE DOMINANT SPICE VARIANT EXPRESSED IN VARIOUS TISSUES. TELOKIN  
 CC HAS BEEN FOUND IN A WIDE VARIETY OF ADULT AND FETAL TISSUES.  
 CC [14]  
 CC PTM: MLCK IS PROBABLY DOWN-REGULATED BY PHOSPHORYLATION.  
 CC [15]  
 CC SIMILARITY: BELONGS TO THE SPR/THP FAMILY OF PROTEIN KINASES.  
 CC [16]  
 CC SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.  
 CC [17]  
 CC SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE III-LIKE DOMAIN.  
 CC [18]  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC [19]  
 CC EMBL: U48959; AAC18423.2;  
 CC EMBL: AF069601; AAD15921.2;

DR EMBL: AF069602; AAD15922.1;  
 DR EMBL: AF069603; AAD15923.1;  
 DR EMBL: AF069604; AAD15924.1;  
 DR EMBL: X85337; CAA53685.1;  
 DR EMBL: AF067771; AAD51380.1;  
 DR EMBL: AF067666; AAD51380.1; JOINED.  
 DR EMBL: AF067667; AAD51380.1; JOINED.  
 DR EMBL: AF067668; AAD51380.1; JOINED.  
 DR EMBL: AF067669; AAD51380.1; JOINED.  
 DR EMBL: AF067670; AAD51380.1; JOINED.  
 DR EMBL: AF067671; AAD51380.1; JOINED.  
 DR EMBL: AF067672; AAD51380.1; JOINED.  
 DR EMBL: AF067673; AAD51380.1; JOINED.  
 DR EMBL: AF067674; AAD51380.1; JOINED.  
 DR EMBL: AF067675; AAD51380.1; JOINED.  
 DR EMBL: AF067676; AAD51380.1; JOINED.  
 DR EMBL: AF067677; AAD51380.1; JOINED.  
 DR EMBL: AF067678; AAD51380.1; JOINED.  
 DR EMBL: AF067679; AAD51380.1; JOINED.  
 DR EMBL: AF067680; AAD51380.1; JOINED.  
 DR EMBL: AF067681; AAD51380.1; JOINED.  
 DR EMBL: AF067682; AAD51380.1; JOINED.  
 DR EMBL: AF067683; AAD51380.1; JOINED.  
 DR EMBL: AF067684; AAD51380.1; JOINED.  
 DR EMBL: AF067685; AAD51380.1; JOINED.  
 DR EMBL: AF067686; AAD51380.1; JOINED.  
 DR EMBL: AF067687; AAD51380.1; JOINED.  
 DR EMBL: AF067688; AAD51380.1; JOINED.  
 DR EMBL: AF067689; AAD51380.1; JOINED.  
 DR EMBL: AF067690; AAD51380.1; JOINED.  
 DR EMBL: AF067691; AAD51380.1; JOINED.  
 DR EMBL: AF067692; AAD51380.1; JOINED.  
 DR EMBL: AF067693; AAD51380.1; JOINED.  
 DR EMBL: AF067694; AAD51380.1; JOINED.  
 DR EMBL: AF067695; AAD51380.1; JOINED.  
 DR EMBL: AF067696; AAD51380.1; JOINED.  
 DR EMBL: AF067697; AAD51380.1; JOINED.  
 DR EMBL: AF067698; AAD51380.1; JOINED.  
 DR EMBL: AF067699; AAD51380.1; JOINED.  
 DR EMBL: AF067700; AAD51380.1; JOINED.  
 DR EMBL: AF067701; AAD51380.1; JOINED.  
 DR EMBL: AF067702; AAD51380.1; JOINED.  
 DR EMBL: AF067703; AAD51380.1; JOINED.  
 DR EMBL: AF067704; AAD51380.1; JOINED.  
 DR EMBL: AF067705; AAD51380.1; JOINED.  
 DR EMBL: AF067706; AAD51380.1; JOINED.  
 DR EMBL: AF067707; AAD51380.1; JOINED.  
 DR EMBL: AF067708; AAD51380.1; JOINED.  
 DR EMBL: AF067709; AAD51380.1; JOINED.  
 DR EMBL: AF067710; AAD51380.1; JOINED.  
 DR EMBL: AF067711; AAD51380.1; JOINED.  
 DR EMBL: AF067712; AAD51380.1; JOINED.  
 DR EMBL: AF067713; AAD51380.1; JOINED.  
 DR EMBL: AF067714; AAD51380.1; JOINED.  
 DR EMBL: AF067715; AAD51380.1; JOINED.  
 DR EMBL: AF067716; AAD51380.1; JOINED.  
 DR EMBL: AF067717; AAD51380.1; JOINED.  
 DR EMBL: AF067718; AAD51380.1; JOINED.  
 DR EMBL: AF067719; AAD51380.1; JOINED.  
 DR EMBL: AF067720; AAD51380.1; JOINED.  
 DR EMBL: AF067721; AAD51380.1; JOINED.  
 DR EMBL: AF067722; AAD51380.1; JOINED.  
 DR EMBL: AF067723; AAD51380.1; JOINED.  
 DR EMBL: AF067724; AAD51380.1; JOINED.  
 DR EMBL: AF067725; AAD51380.1; JOINED.  
 DR EMBL: AF067726; AAD51380.1; JOINED.  
 DR EMBL: AF067727; AAD51380.1; JOINED.  
 DR EMBL: AF067728; AAD51380.1; JOINED.  
 DR EMBL: AF067729; AAD51380.1; JOINED.  
 DR EMBL: AF067730; AAD51380.1; JOINED.  
 DR EMBL: AF067731; AAD51380.1; JOINED.  
 DR EMBL: AF067732; AAD51380.1; JOINED.  
 DR EMBL: AF067733; AAD51380.1; JOINED.  
 DR EMBL: AF067734; AAD51380.1; JOINED.  
 DR EMBL: AF067735; AAD51380.1; JOINED.  
 DR EMBL: AF067736; AAD51380.1; JOINED.  
 DR EMBL: AF067737; AAD51380.1; JOINED.  
 DR EMBL: AF067738; AAD51380.1; JOINED.  
 DR EMBL: AF067739; AAD51380.1; JOINED.  
 DR EMBL: AF067740; AAD51380.1; JOINED.  
 DR EMBL: AF067741; AAD51380.1; JOINED.  
 DR EMBL: AF067742; AAD51380.1; JOINED.  
 DR EMBL: AF067743; AAD51380.1; JOINED.  
 DR EMBL: AF067744; AAD51380.1; JOINED.  
 DR EMBL: AF067745; AAD51380.1; JOINED.  
 DR EMBL: AF067746; AAD51380.1; JOINED.  
 DR EMBL: AF067747; AAD51380.1; JOINED.  
 DR EMBL: AF067748; AAD51380.1; JOINED.  
 DR EMBL: AF067749; AAD51380.1; JOINED.  
 DR EMBL: AF067750; AAD51380.1; JOINED.  
 DR EMBL: AF067751; AAD51380.1; JOINED.  
 DR EMBL: AF067752; AAD51380.1; JOINED.  
 DR EMBL: AF067753; AAD51380.1; JOINED.  
 DR EMBL: AF067754; AAD51380.1; JOINED.  
 DR EMBL: AF067755; AAD51380.1; JOINED.  
 DR EMBL: AF067756; AAD51380.1; JOINED.  
 DR EMBL: AF067757; AAD51380.1; JOINED.  
 DR EMBL: AF067758; AAD51380.1; JOINED.  
 DR EMBL: AF067759; AAD51380.1; JOINED.  
 DR EMBL: AF067760; AAD51380.1; JOINED.  
 DR EMBL: AF067761; AAD51380.1; JOINED.  
 DR EMBL: AF067762; AAD51380.1; JOINED.  
 DR EMBL: AF067763; AAD51380.1; JOINED.  
 DR EMBL: AF067764; AAD51380.1; JOINED.  
 DR EMBL: AF067765; AAD51380.1; JOINED.  
 DR EMBL: AF067766; AAD51380.1; JOINED.  
 DR EMBL: AF067767; AAD51380.1; JOINED.  
 DR EMBL: AF067768; AAD51380.1; JOINED.  
 DR EMBL: AF067769; AAD51380.1; JOINED.  
 DR EMBL: AF067770; AAD51380.1; JOINED.  
 DR EMBL: AF067771; AAD51380.1; JOINED.  
 DR EMBL: AF067772; AAD51380.1; JOINED.  
 DR EMBL: AF067773; AAD51380.1; JOINED.  
 DR EMBL: AF067774; AAD51380.1; JOINED.  
 DR EMBL: AF067775; AAD51380.1; JOINED.  
 DR EMBL: AF067776; AAD51380.1; JOINED.  
 DR EMBL: AF067777; AAD51380.1; JOINED.  
 DR EMBL: AF067778; AAD51380.1; JOINED.  
 DR EMBL: AF067779; AAD51380.1; JOINED.  
 DR EMBL: AF067780; AAD51380.1; JOINED.  
 DR EMBL: AF067781; AAD51380.1; JOINED.  
 DR EMBL: AF067782; AAD51380.1; JOINED.  
 DR EMBL: AF067783; AAD51380.1; JOINED.  
 DR EMBL: AF067784; AAD51380.1; JOINED.  
 DR EMBL: AF067785; AAD51380.1; JOINED.  
 DR EMBL: AF067786; AAD51380.1; JOINED.  
 DR EMBL: AF067787; AAD51380.1; JOINED.  
 DR EMBL: AF067788; AAD51380.1; JOINED.  
 DR EMBL: AF067789; AAD51380.1; JOINED.  
 DR EMBL: AF067790; AAD51380.1; JOINED.  
 DR EMBL: AF067791; AAD51380.1; JOINED.  
 DR EMBL: AF067792; AAD51380.1; JOINED.  
 DR EMBL: AF067793; AAD51380.1; JOINED.  
 DR EMBL: AF067794; AAD51380.1; JOINED.  
 DR EMBL: AF067795; AAD51380.1; JOINED.  
 DR EMBL: AF067796; AAD51380.1; JOINED.  
 DR EMBL: AF067797; AAD51380.1; JOINED.  
 DR EMBL: AF067798; AAD51380.1; JOINED.  
 DR EMBL: AF067799; AAD51380.1; JOINED.  
 DR EMBL: AF067800; AAD51380.1; JOINED.  
 DR EMBL: AF067801; AAD51380.1; JOINED.  
 DR EMBL: AF067802; AAD51380.1; JOINED.  
 DR EMBL: AF067803; AAD51380.1; JOINED.  
 DR EMBL: AF067804; AAD51380.1; JOINED.  
 DR EMBL: AF067805; AAD51380.1; JOINED.  
 DR EMBL: AF067806; AAD51380.1; JOINED.  
 DR EMBL: AF067807; AAD51380.1; JOINED.  
 DR EMBL: AF067808; AAD51380.1; JOINED.  
 DR EMBL: AF067809; AAD51380.1; JOINED.  
 DR EMBL: AF067810; AAD51380.1; JOINED.  
 DR EMBL: AF067811; AAD51380.1; JOINED.  
 DR EMBL: AF067812; AAD51380.1; JOINED.  
 DR EMBL: AF067813; AAD51380.1; JOINED.  
 DR EMBL: AF067814; AAD51380.1; JOINED.  
 DR EMBL: AF067815; AAD51380.1; JOINED.  
 DR EMBL: AF067816; AAD51380.1; JOINED.  
 DR EMBL: AF067817; AAD51380.1; JOINED.  
 DR EMBL: AF067818; AAD51380.1; JOINED.  
 DR EMBL: AF067819; AAD51380.1; JOINED.  
 DR EMBL: AF067820; AAD51380.1; JOINED.  
 DR EMBL: AF067821; AAD51380.1; JOINED.  
 DR EMBL: AF067822; AAD51380.1; JOINED.  
 DR EMBL: AF067823; AAD51380.1; JOINED.  
 DR EMBL: AF067824; AAD51380.1; JOINED.  
 DR EMBL: AF067825; AAD51380.1; JOINED.  
 DR EMBL: AF067826; AAD51380.1; JOINED.  
 DR EMBL: AF067827; AAD51380.1; JOINED.  
 DR EMBL: AF067828; AAD51380.1; JOINED.  
 DR EMBL: AF067829; AAD51380.1; JOINED.  
 DR EMBL: AF067830; AAD51380.1; JOINED.  
 DR EMBL: AF067831; AAD51380.1; JOINED.  
 DR EMBL: AF067832; AAD51380.1; JOINED.  
 DR EMBL: AF067833; AAD51380.1; JOINED.  
 DR EMBL: AF067834; AAD51380.1; JOINED.  
 DR EMBL: AF067835; AAD51380.1; JOINED.  
 DR EMBL: AF067836; AAD51380.1; JOINED.  
 DR EMBL: AF067837; AAD51380.1; JOINED.  
 DR EMBL: AF067838; AAD51380.1; JOINED.  
 DR EMBL: AF067839; AAD51380.1; JOINED.  
 DR EMBL: AF067840; AAD51380.1; JOINED.  
 DR EMBL: AF067841; AAD51380.1; JOINED.  
 DR EMBL: AF067842; AAD51380.1; JOINED.  
 DR EMBL: AF067843; AAD51380.1; JOINED.  
 DR EMBL: AF067844; AAD51380.1; JOINED.  
 DR EMBL: AF067845; AAD51380.1; JOINED.  
 DR EMBL: AF067846; AAD51380.1; JOINED.  
 DR EMBL: AF067847; AAD51380.1; JOINED.  
 DR EMBL: AF067848; AAD51380.1; JOINED.  
 DR EMBL: AF067849; AAD51380.1; JOINED.  
 DR EMBL: AF067850; AAD51380.1; JOINED.  
 DR EMBL: AF067851; AAD51380.1; JOINED.  
 DR EMBL: AF067852; AAD51380.1; JOINED.  
 DR EMBL: AF067853; AAD51380.1; JOINED.  
 DR EMBL: AF067854; AAD51380.1; JOINED.  
 DR EMBL: AF067855; AAD51380.1; JOINED.  
 DR EMBL: AF067856; AAD51380.1; JOINED.  
 DR EMBL: AF067857; AAD51380.1; JOINED.  
 DR EMBL: AF067858; AAD51380.1; JOINED.  
 DR EMBL: AF067859; AAD51380.1; JOINED.  
 DR EMBL: AF067860; AAD51380.1; JOINED.  
 DR EMBL: AF067861; AAD51380.1; JOINED.  
 DR EMBL: AF067862; AAD51380.1; JOINED.  
 DR EMBL: AF067863; AAD51380.1; JOINED.  
 DR EMBL: AF067864; AAD51380.1; JOINED.  
 DR EMBL: AF067865; AAD51380.1; JOINED.  
 DR EMBL: AF067866; AAD51380.1; JOINED.  
 DR EMBL: AF067867; AAD51380.1; JOINED.  
 DR EMBL: AF067868; AAD51380.1; JOINED.  
 DR EMBL: AF067869; AAD51380.1; JOINED.  
 DR EMBL: AF067870; AAD51380.1; JOINED.  
 DR EMBL: AF067871; AAD51380.1; JOINED.  
 DR EMBL: AF067872; AAD51380.1; JOINED.  
 DR EMBL: AF067873; AAD51380.1; JOINED.  
 DR EMBL: AF067874; AAD51380.1; JOINED.  
 DR EMBL: AF067875; AAD51380.1; JOINED.  
 DR EMBL: AF067876; AAD51380.1; JOINED.  
 DR EMBL: AF067877; AAD51380.1; JOINED.  
 DR EMBL: AF067878; AAD51380.1; JOINED.  
 DR EMBL: AF067879; AAD51380.1; JOINED.  
 DR EMBL: AF067880; AAD51380.1; JOINED.  
 DR EMBL: AF067881; AAD51380.1; JOINED.  
 DR EMBL: AF067882; AAD51380.1; JOINED.  
 DR EMBL: AF067883; AAD51380.1; JOINED.  
 DR EMBL: AF067884; AAD51380.1; JOINED.  
 DR EMBL: AF067885; AAD51380.1; JOINED.  
 DR EMBL: AF067886; AAD51380.1; JOINED.  
 DR EMBL: AF067887; AAD51380.1; JOINED.  
 DR EMBL: AF067888; AAD51380.1; JOINED.  
 DR EMBL: AF067889; AAD51380.1; JOINED.  
 DR EMBL: AF067890; AAD51380.1; JOINED.  
 DR EMBL: AF067891; AAD51380.1; JOINED.  
 DR EMBL: AF067892; AAD51380.1; JOINED.  
 DR EMBL: AF067893; AAD51380.1; JOINED.  
 DR EMBL: AF067894; AAD51380.1; JOINED.  
 DR EMBL: AF067895; AAD51380.1; JOINED.  
 DR EMBL: AF067896; AAD51380.1; JOINED.  
 DR EMBL: AF067897; AAD51380.1; JOINED.  
 DR EMBL: AF067898; AAD51380.1; JOINED.  
 DR EMBL: AF067899; AAD51380.1; JOINED.  
 DR EMBL: AF067900; AAD51380.1; JOINED.  
 DR EMBL: AF067901; AAD51380.1; JOINED.  
 DR EMBL: AF067902; AAD51380.1; JOINED.  
 DR EMBL: AF067903; AAD51380.1; JOINED.  
 DR EMBL: AF067904; AAD51380.1; JOINED.  
 DR EMBL: AF067905; AAD51380.1; JOINED.  
 DR EMBL: AF067906; AAD51380.1; JOINED.  
 DR EMBL: AF067907; AAD51380.1; JOINED.  
 DR EMBL: AF067908; AAD51380.1; JOINED.  
 DR EMBL: AF067909; AAD51380.1; JOINED.  
 DR EMBL: AF067910; AAD51380.1; JOINED.  
 DR EMBL: AF067911; AAD51380.1; JOINED.  
 DR EMBL: AF067912; AAD51380.1; JOINED.  
 DR EMBL: AF067913; AAD51380.1; JOINED.  
 DR EMBL: AF067914; AAD51380.1; JOINED.  
 DR EMBL: AF067915; AAD51380.1; JOINED.  
 DR EMBL: AF067916; AAD51380.1; JOINED.  
 DR EMBL: AF067917; AAD51380.1; JOINED.  
 DR EMBL: AF067918; AAD51380.1; JOINED.  
 DR EMBL: AF067919; AAD51380.1; JOINED.  
 DR EMBL: AF067920; AAD51380.1; JOINED.  
 DR EMBL: AF067921; AAD51380.1; JOINED.  
 DR EMBL: AF067922; AAD51380.1; JOINED.  
 DR EMBL: AF067923; AAD51380.1; JOINED.  
 DR EMBL: AF067924; AAD51380.1; JOINED.  
 DR EMBL: AF067925; AAD51380.1; JOINED.  
 DR EMBL: AF067926; AAD51380.1; JOINED.  
 DR EMBL: AF067927; AAD51380.1; JOINED.  
 DR EMBL: AF067928; AAD51380.1; JOINED.  
 DR EMBL: AF067929; AAD51380.1; JOINED.  
 DR EMBL: AF067930; AAD51380.1; JOINED.  
 DR EMBL: AF067931; AAD51380.1; JOINED.  
 DR EMBL: AF067932; AAD51380.1; JOINED.  
 DR EMBL: AF067933; AAD51380.1; JOINED.  
 DR EMBL: AF067934; AAD51380.1; JOINED.  
 DR EMBL: AF067935; AAD51380.1; JOINED.  
 DR EMBL: AF067936; AAD51380.1; JOINED.  
 DR EMBL: AF067937; AAD51380.1; JOINED.  
 DR EMBL: AF067938; AAD51380.1; JOINED.  
 DR EMBL: AF067939; AAD51380.1; JOINED.  
 DR EMBL: AF067940; AAD51380.1; JOINED.  
 DR EMBL: AF067941; AAD51380.1; JOINED.  
 DR EMBL: AF067942; AAD51380.1; JOINED.  
 DR EMBL: AF067943; AAD51380.1; JOINED.  
 DR EMBL: AF067944; AAD51380.1; JOINED.  
 DR EMBL: AF067945; AAD51380.1; JOINED.  
 DR EMBL: AF067946; AAD51380.1; JOINED.  
 DR EMBL: AF067947; AAD51380.1; JOINED.  
 DR EMBL: AF067948; AAD51380.1; JOINED.  
 DR EMBL: AF067949; AAD51380.1; JOINED.  
 DR EMBL: AF067950; AAD51380.1; JOINED.  
 DR EMBL: AF067951; AAD51380.1; JOINED.  
 DR EMBL: AF067952; AAD51380.1; JOINED.  
 DR EMBL: AF067953; AAD51380.1; JOINED.  
 DR EMBL: AF067954; AAD51380.1; JOINED.  
 DR EMBL: AF067955; AAD51380.1; JOINED.  
 DR EMBL: AF067956; AAD51380.1; JOINED.  
 DR EMBL: AF067957; AAD51380.1; JOINED.  
 DR EMBL: AF067958; AAD51380.1; JOINED.  
 DR EMBL: AF067959; AAD51380.1; JOINED.  
 DR EMBL: AF067960; AAD51380.1; JOINED.  
 DR EMBL: AF067961; AAD51380.1; JOINED.  
 DR EMBL: AF067962; AAD51380.1; JOINED.  
 DR EMBL: AF067963; AAD51380.1; JOINED.  
 DR EMBL: AF067964; AAD51380.1; JOINED.  
 DR EMBL: AF067965; AAD51380.1; JOINED.  
 DR EMBL: AF067966; AAD51380.1; JOINED.  
 DR EMBL: AF067967; AAD51380.1; JOINED.  
 DR EMBL: AF067968; AAD51380.1; JOINED.  
 DR EMBL: AF067969; AAD51380.1; JOINED.  
 DR EMBL: AF067970; AAD51380.1; JOINED.  
 DR EMBL: AF067971; AAD51380.1; JOINED.  
 DR EMBL: AF067972; AAD51380.1; JOINED.  
 DR EMBL: AF067973; AAD51380.1; JOINED.  
 DR EMBL: AF067974; AAD51380.1; JOINED.  
 DR EMBL: AF067975; AAD51380.1; JOINED.  
 DR EMBL: AF067976; AAD51380.1; JOINED.  
 DR EMBL: AF067977; AAD51380.1; JOINED.  
 DR EMBL: AF067978; AAD51380.1; JOINED.  
 DR EMBL: AF067979; AAD51380.1; JOINED.  
 DR EMBL: AF067980; AAD51380.1; JOINED.  
 DR EMBL: AF067981; AAD51380.1; JOINED.  
 DR EMBL: AF067982; AAD51380.1; JOINED.  
 DR EMBL: AF067983; AAD51380.1; JOINED.  
 DR EMBL: AF067984; AAD51380.1; JOINED.  
 DR EMBL: AF067985; AAD51380.1; JOINED.  
 DR EMBL: AF067986; AAD51380.1; JOINED.  
 DR EMBL: AF067987; AAD51380.1; JOINED.  
 DR EMBL: AF067988; AAD51380.1; JOINED.  
 DR EMBL: AF067989; AAD51380.1; JOINED.  
 DR EMBL: AF067990; AAD51380.1; JOINED.  
 DR EMBL: AF067991; AAD51380.1; JOINED.  
 DR EMBL: AF067992; AAD51380.1; JOINED.  
 DR EMBL: AF067993; AAD51380.1; JOINED.  
 DR EMBL: AF067994; AAD51380.1; JOINED.  
 DR EMBL: AF067995; AAD51380.1; JOINED.  
 DR EMBL: AF067996; AAD51380.1; JOINED.  
 DR EMBL: AF067997; AAD51380.1; JOINED.  
 DR EMBL: AF067998; AAD51380.1; JOINED.  
 DR EMBL: AF067999; AAD51380.1; JOINED.  
 DR EMBL: AF068000; AAD51380.1; JOINED.  
 DR EMBL: AF068001; AAD51380.1; JOINED.  
 DR EMBL: AF068002; AAD51380.1; JOINED.  
 DR EMBL: AF068003; AAD51380.1; JOINED.  
 DR EMBL: AF068004; AAD51380.1; JOINED.  
 DR EMBL: AF068005; AAD51380.1; JOINED.  
 DR EMBL: AF068006; AAD51380.1; JOINED.  
 DR EMBL: AF068007; AAD51380.1; JOINED.  
 DR EMBL: AF068008; AAD51380.1; JOINED.  
 DR EMBL: AF068009; AAD51380.1; JOINED.  
 DR EMBL: AF068010; AAD51380.1; JOINED.  
 DR EMBL: AF068011; AAD51380.1; JOINED.  
 DR EMBL: AF068012; AAD51380.1; JOINED.  
 DR EMBL: AF068013; AAD51380.1; JOINED.  
 DR





QY 110 ELTAWENK--VGLVWSESGYFAGTILSW--HILGKPLVPNEKGVAVKEGIRPPTSLFT 165  
 DB 158 PAKGKDKVAVIACTISANKCKPVSVMWTHLKEG-----AEYGEIPNP-NGIVT 206  
 QY 166 LQSELMVTPARAGDPPTESNSFSLPHRALRTAPIQVRYWELVPLVEVLVYVPE-EG 244  
 DB 207 VLSKRYLVPKR-EDHQRSLACIVNVHMRKESLILNVQ---YEP-----EVTIESPDG 256  
 QY 225 GAVAPGCTVTHGFEVPAQ-PSQIHWKKGCVPIP---LPPSPVILIPICQDQCTYSC 279  
 DB 257 NWTQLQMDVKLLKALANFPALEYHWTLNLSLPKAVEA-NKTLFFPAPINTSMASTYIP 316  
 QY 280 VATHSSHFPESKA--VSTIIIE-----PREEPTASVQ-----GSLPPTIALAUGILR 327  
 DB 317 EAIN-----PIGTHSQGVENITFEPTYPSPDPRGRRACQVPTAIIICGVVGSILLVIFVVG 372  
 QY 328 GAGTA 332  
 DB 373 GIVVA 377

RESULT 15  
 PGDM\_MOUSE STANDARD: PRT: 3707 AA  
 ID PGDM\_MOUSE  
 AC Q05793;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, last sequence update)  
 DT 15-JUN-2002 (Rel. 41, last annotation update)  
 DE Basament membrane-specific heparan sulfate proteoglycan core  
 DE protein precursor (HSPG) (Perlecan) (PDC).  
 GN HSPG2.  
 OS Mus musculus (Mouse)  
 CC Fukaya, M., McTague, C., Choudhary, V., Horigan, E., Sasaki, M.,  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N A  
 RC HSSSE-Melanoma;  
 RX MEDLINE=92078153; PubMed=1744097;  
 RA Noonan, D.M., Fulle, A., Valente, P., Cai, S., Horigan, E., Sasaki, M.,  
 RA Yamada, Y., Hassell, J.P.;  
 RT "The complete sequence of perlecan, a basement membrane heparan  
 RT sulfate proteoglycan, reveals extensive similarity with laminin A  
 RT chain, low density lipoprotein receptor, and the neural cell adhesion  
 RT molecule.";  
 RL J. Biol. Chem. 264:22039-22047(1991)  
 RN [2]  
 RP SEQUENCE OF 940-1601 AND 1870-2600 FROM N A, AND PARTIAL SEQUENCE.  
 RX MEDLINE=89034110, PubMed=292709,  
 RA Noonan, D.M., Horigan, E.A., Ledbetter, S.R., Vogel, C., Sasaki, M.,  
 RA Yamada, Y., Hassell, J.P.;  
 RT "Identification of cDNA clones encoding different domains of the  
 RT basement membrane heparan sulfate proteoglycan.";  
 RL J. Biol. Chem. 263:16379-16387(1988).  
 CC -1- FUNCTION: This protein is an integral component of basement  
 CC membranes. It is responsible for the fixed negative electrostatic  
 CC charge and is involved in the charge-selective ultrafiltration  
 CC properties. It serves as an attachment substrate for cells.  
 CC -1- SUBUNIT: Purified perlecan has a strong tendency to aggregate in  
 CC dimers or stellate structures. It interacts with other basement  
 CC membrane components such as laminin, proteoglycan and collagen type-  
 CC IV.  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES.  
 CC -1- PTC: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED  
 CC AND O-LINKED GLYCOSAMINOGLYCAN.  
 CC -1- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 10-15 LAMININ EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 LAMININ DOMAINS IV.  
 CC -1- SIMILARITY: CONTAINS 15 LAMININ EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 LAMININ EGF-LIKE DOMAINS  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN

-1- SIMILARITY: CONTAINS 1 SEA DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M77174; AAA39911.1; -;  
 CC EMBL: J04954; AAA39991.1; -;  
 CC EMBL: J04055; AAA39912.1; -;  
 CC HSP: P01130; IAJJ.  
 CC MGD: MGI:96257; Hsp92.  
 CC InterPro: IPR000561; EGF-like.  
 CC InterPro: IPR000742; EGF-2.  
 CC InterPro: IPR001438; EGF-II.  
 CC InterPro: IPR003006; IG\_MHC.  
 CC InterPro: IPR003598; IG\_C2.  
 CC InterPro: IPR002172; LDL\_recept\_A.  
 CC InterPro: IPR000034; Laminin\_B.  
 CC InterPro: IPR002049; Laminin\_EGF.  
 CC InterPro: IPR001791; Laminin\_G.  
 CC InterPro: IPR000082; SEA\_domain.  
 CC Pfam: PF00008; EGF; 4.  
 CC Pfam: PF00047; Ig; 15.  
 CC Pfam: PF00052; laminin\_B; 3.  
 CC Pfam: PF00053; laminin\_EGF; 8.  
 CC Pfam: PF00054; laminin\_G; 3.  
 CC Pfam: PF00057; ldl\_recept\_a; 4.  
 CC Pfam: PF01390; SEA; 1.  
 CC PRINTS: PF00010; EGFRII\_OOD.  
 CC PROSITE: PS003031; Laminin\_B; 3.  
 CC SMART: SM00180; EGF\_Lam; 7.  
 CC SMART: SM00001; EGF\_Like; 6.  
 CC SMART: SM00408; IGC2; 14.  
 CC SMART: SM00192; LDLa; 4.  
 CC SMART: SM00281; lamr; 3.  
 CC SMART: SM00282; lamc; 3.  
 CC SMART: SM00200; SEA; 1.  
 CC PROSITE: PS00022; EGF\_1; 8.  
 CC PROSITE: PS01186; EGF\_2; 5.  
 CC PROSITE: PS01248; LAMININ\_TYPE\_EGF; 11.  
 CC PROSITE: PS00025; LAM\_5\_DOMAIN; 3.  
 CC PROSITE: PS01209; LDLRA\_1; 4.  
 CC PROSITE: PS00068; LDLRA\_2; 4.  
 CC PROSITE: PS00024; SEA; 1.  
 CC Signal: Basement membrane; Proteoglycan; Repeat; Glycosylated;  
 CC Heparan sulfate; Laminin; EGF-like domain; Immunoglobulin domain;  
 CC Extracellular matrix; EGF-like domain.  
 CC SIGNAL: 1 21  
 CC CHAIN: 22 3707  
 CC  
 CC DOMAIN 80 194  
 CC DOMAIN 195 234  
 CC DOMAIN 281 319  
 CC DOMAIN 320 359  
 CC DOMAIN 360 403  
 CC DOMAIN 404 504  
 CC DOMAIN 521 530  
 CC DOMAIN 531 730  
 CC DOMAIN 731 763  
 CC DOMAIN 764 813  
 CC DOMAIN 814 871  
 CC DOMAIN 879 923  
 CC DOMAIN 924 933  
 CC DOMAIN 934 1125  
 CC DOMAIN 1126 1158  
 CC DOMAIN 1159 1208  
 CC DOMAIN 1209 1255  
 CC DOMAIN 1275 1324  
 CC DOMAIN 1325 1334  
 CC  
 CC LDL-RECEPTOR CLASS A 1  
 CC LDL-RECEPTOR CLASS A 2  
 CC LDL-RECEPTOR CLASS A 3  
 CC LDL-RECEPTOR CLASS A 4  
 CC IG-LIKE C2-TYPE DOMAIN 1  
 CC LAMININ EGF-LIKE 1 (N-TERMINAL)  
 CC LAMININ DOMAIN IV 1 (DOMAIN III A).  
 CC LAMININ EGF-LIKE 1 (C-TERMINAL).  
 CC LAMININ EGF-LIKE 2.  
 CC LAMININ EGF-LIKE 3.  
 CC LAMININ EGF-LIKE 4 (INCOMPLETE).  
 CC LAMININ EGF-LIKE 5 (N-TERMINAL).  
 CC LAMININ EGF-LIKE 5 (C-TERMINAL).  
 CC LAMININ EGF-LIKE 5 (C-TERMINAL).  
 CC LAMININ EGF-LIKE 6.  
 CC LAMININ EGF-LIKE 7.  
 CC LAMININ EGF-LIKE 8.  
 CC LAMININ EGF-LIKE 9 (N-TERMINAL).  
 CC LAMININ EGF-LIKE 9 (N-TERMINAL).

FL	DOMAIN	1435	1529	AMININ DOMAIN IV 3 (DOMAIN 111 C)
FT	DOMAIN	1530	1562	AMININ EGF-LIKE 9 (C TERMINAL)
FT	DOMAIN	1563	1612	AMININ EGF-LIKE 10
FT	DOMAIN	1614	1670	AMININ EGF-LIKE 11
FT	DOMAIN	1677	1771	IG LIKE C2 TYPE DOMAIN 2
FL	DOMAIN	1772	1865	IG LIKE C2 TYPE DOMAIN 4
FL	DOMAIN	1866	1954	IG LIKE C2 TYPE DOMAIN 4
FL	DOMAIN	1955	2049	IG LIKE C2 TYPE DOMAIN 5
FT	DOMAIN	2050	2148	IG LIKE C2 TYPE DOMAIN 6
FT	DOMAIN	2149	2244	IG LIKE C2 TYPE DOMAIN 7
FT	DOMAIN	2245	2343	IG LIKE C2 TYPE DOMAIN 8
FL	DOMAIN	2344	2436	IG LIKE C2 TYPE DOMAIN 9
FL	DOMAIN	2437	2532	IG LIKE C2 TYPE DOMAIN 10
FT	DOMAIN	2533	2619	IG LIKE C2 TYPE DOMAIN 11
FT	DOMAIN	2620	2726	IG LIKE C2 TYPE DOMAIN 12
FT	DOMAIN	2727	2809	IG LIKE C2 TYPE DOMAIN 13
FT	DOMAIN	2810	2895	IG LIKE C2 TYPE DOMAIN 14
FL	DOMAIN	2896	2980	IG LIKE C2 TYPE DOMAIN 15
FL	DOMAIN	2984	3162	IG LIKE C2 TYPE DOMAIN 15
FL	DOMAIN	3163	3241	EGF-LIKE
FL	DOMAIN	3245	3425	AMININ G-LIKE 2
FT	DOMAIN	3426	3705	AMININ G-LIKE 3
FT	DOMAIN	3706	3775	HEPARAN SULFATE (POTENTIAL)
FL	SITE	65	67	HEPARAN SULFATE (POTENTIAL)
FL	SITE	71	73	HEPARAN SULFATE (POTENTIAL)
FL	SITE	76	78	HEPARAN SULFATE (POTENTIAL)
FL	SITE	76	78	MEDIATES MOTOR NEURON ATTACHMENT (POTENTIAL)
FT	DISULFID	3615	3617	BY SIMILARITY
FT	DISULFID	199	212	BY SIMILARITY
FT	DISULFID	206	225	BY SIMILARITY
FL	DISULFID	219	234	BY SIMILARITY
FL	DISULFID	285	297	BY SIMILARITY
FL	DISULFID	292	310	BY SIMILARITY
FL	DISULFID	404	419	BY SIMILARITY
FL	DISULFID	425	437	BY SIMILARITY
FL	DISULFID	442	450	BY SIMILARITY
FL	DISULFID	444	459	BY SIMILARITY
FL	DISULFID	468	481	BY SIMILARITY
FL	DISULFID	475	504	BY SIMILARITY
FL	DISULFID	488	493	BY SIMILARITY
FL	DISULFID	498	479	BY SIMILARITY
FL	DISULFID	764	773	BY SIMILARITY
FL	DISULFID	766	780	BY SIMILARITY
FT	DISULFID	783	792	BY SIMILARITY
FL	DISULFID	795	811	BY SIMILARITY
FL	DISULFID	814	829	BY SIMILARITY
FL	DISULFID	816	839	BY SIMILARITY
FL	DISULFID	842	851	BY SIMILARITY
FL	DISULFID	854	869	BY SIMILARITY
FL	DISULFID	1159	1168	BY SIMILARITY
FL	DISULFID	1161	1175	BY SIMILARITY
FT	DISULFID	1178	1187	BY SIMILARITY
FT	DISULFID	1190	1206	BY SIMILARITY
FT	DISULFID	1209	1224	BY SIMILARITY
FL	DISULFID	1211	1234	BY SIMILARITY
FL	DISULFID	1237	1246	BY SIMILARITY
FL	DISULFID	1249	1263	BY SIMILARITY
FT	DISULFID	1275	1287	BY SIMILARITY
FT	DISULFID	1293	1293	BY SIMILARITY
FT	DISULFID	1307	1322	BY SIMILARITY
FL	DISULFID	1563	1572	BY SIMILARITY
FL	DISULFID	1565	1579	BY SIMILARITY
FT	DISULFID	1582	1591	BY SIMILARITY
FT	DISULFID	1594	1610	BY SIMILARITY
FL	DISULFID	1613	1628	BY SIMILARITY
FL	DISULFID	1615	1638	BY SIMILARITY
FL	DISULFID	1641	1650	BY SIMILARITY
FL	DISULFID	1653	1668	BY SIMILARITY
FL	DISULFID	1792	1839	BY SIMILARITY
FL	DISULFID	1886	1932	BY SIMILARITY
FL	DISULFID	1976	2021	BY SIMILARITY
FL	DISULFID	2073	2118	BY SIMILARITY
FL	DISULFID	2170	2215	BY SIMILARITY

FT	DISULF ID	2268	2413	BY SIMILARITY		
FT	DISULF ID	2365	2413	BY SIMILARITY		
FT	DISULF ID	2456	2506	BY SIMILARITY		
FT	DISULF ID	2554	2599	BY SIMILARITY		
FT	DISULF ID	2641	2686	BY SIMILARITY		
Query Match					10.5%; Score (85); Len 1; Locals (202);	
Best Local Similarity					25.7%; Pred. No. 0.00046;	
Matches					81; Conservative 43; Mismatches 121; Indels 80; Gaps 16;	
QY	5 TARIQELVLKQKAIKKPKPQDLEKNTGRTKAMKVLSPQSGIWNQVAKVLLNSLFL					64
DB	2354 TVQQQQAAPKCLHREKANGVKKRTRQELLE					2399
QY	64 LEAVATLQGTGTFVYKAMNPK FFKSNFVR					118
DB	2406 TVAPGATMEPTACVASKNYVMAQSVNLNVSRPTVAVLEP FVHKWQKDTLL					2454
QY	119 VGIQVSEGSYPACTLSWHLDDKPLVNRKGVSVKQTRRRHLEGLFQSLFMVLTARAG					178
DB	2455 ECLSSSE FSSSPFW					2499
QY	179 DRRPTGSSPSGCLPHRPAIRTAIPQPKWEPVPEVLLVVEPESAVAP					229
DB	2500 D-ACTVYVQ					2539
QY	230 ----- GRTVTLTCEVPAQSSPQHWKDAVPLP					279
DB	2546 SELTLEAHITATLHGSGATGTFPTTHWSKRAPIHWQHTEINTLVTFVAGQSSQYV					2599
QY	280 VAHSSHQIQPSRAV					294
DB	2600 NATNSA GHTEATV					2613

GenBank version 5.1.6  
Copyright (c) 1993 - 2003 Computer Ltd.

OM protein - protein search using sw model

Run on: May 30, 2003, 16:47:46; Search time 77.018 seconds  
(without alignments)  
888,203 Million cell updates/sec

Title: US-09-872-185B-2  
Pertinent scores: 1766  
Sequence: 1 AUNIIARIGEPVLAK:KGAP.....ASQGLIALALGHHIDATA 342

Scoring table: BLQSOM62  
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 20047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 20047115

Post processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SPREMBL-21:

- 1: sp\_archaea:
- 2: sp\_bacteria:
- 3: sp\_fungi:
- 4: sp\_human:
- 5: sp\_invertebrate:
- 6: sp\_mammal:
- 7: sp\_mhc:
- 8: sp\_orqanelle:
- 9: sp\_phage:
- 10: sp\_plant:
- 11: sp\_rodent:
- 12: sp\_virus:
- 13: sp\_vertebrate:
- 14: sp\_unclassified:
- 15: sp\_virus:
- 16: sp\_bacteriophage:
- 17: sp\_archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1381	78.2	402	11	Q35444 mus musculus
2	224.5	12.7	583	11	Q8R270 mus musculus
3	220	12.5	521	6	Q46651 cryotolagus
4	219.5	12.4	646	4	Q95812 homo sapien
5	218.5	12.4	646	4	Q95812 homo sapien
6	217	12.3	1344	11	Q95234 mus musculus
7	214	12.1	1341	4	Q95234 mus musculus
8	213	12.1	583	11	Q35112 mus musculus
9	210.5	11.9	340	13	Q95242 gallus galli
10	210	11.9	1496	4	Q95242 gallus galli
11	209	11.8	1612	11	Q95242 gallus galli
12	209	11.8	1651	4	Q95242 gallus galli
13	208	11.8	1043	5	Q95242 gallus galli
14	204.5	11.6	583	6	Q95813 mus musculus
15	204	11.6	521	6	Q46634 canis fami
16	204	11.6	5636	4	Q95813 mus musculus

17	203	11.5	1033	5	Q9V643 drosophila
18	203	11.5	1651	11	Q55005 rattus norv
19	201	11.4	743	11	Q70246 mus musculus
20	201	11.4	1389	13	Q90259 brachydanio
21	201	11.4	1419	13	Q98583 brachydanio
22	198.5	11.3	606	11	Q46558 mus musculus
23	198.5	11.2	548	11	Q95812 rattus norv
24	198	11.2	4770	4	Q9H3V5 homo sapien
25	196.5	11.1	1614	13	Q8U0V7 xenopus lae
26	196	11.1	333	13	Q95241 gallus galli
27	195.5	11.1	1380	4	Q9HCK4 homo sapien
28	195	11.0	1102	11	Q923W7 mus musculus
29	194.5	11.0	584	13	Q98921 gallus galli
30	194.5	11.0	584	13	Q98921 gallus galli
31	194.5	11.0	584	13	Q98921 gallus galli
32	194.5	11.0	584	13	Q98921 gallus galli
33	194.5	11.0	626	13	Q98921 gallus galli
34	192.5	10.9	1032	13	Q86VD6 mus musculus
35	192.5	10.9	1032	13	Q86VD6 mus musculus
36	192.5	10.9	1032	13	Q86VD6 mus musculus
37	191	10.8	1675	13	Q98SW4 brachydanio
38	189.5	10.7	1273	5	Q44928 caenorhabdi
39	186.5	10.6	1513	13	Q90270 brachydanio
40	186.5	10.6	1513	13	Q90270 brachydanio
41	185.5	10.5	444	4	Q95121 homo sapien
42	185.5	10.5	521	11	Q41352 mus musculus
43	185.5	10.5	939	5	Q957X6 drosophila
44	184.5	10.4	344	11	Q95PJ0 mus musculus
45	183.5	10.4	1114	4	Q9BWV1 homo sapien

## ALIGNMENTS

RESULT 1

ID	Q35444	PRELIMINARY:	PRT:	402 AA.
AC	Q35444			
DT	01-JAN-1998 (TREMBL)			05, Created)
DT	01-JAN-1998 (TREMBL)			05, Last sequence update)
DT	01-DEC-2001 (TREMBL)			19, Last annotation update)
DE	RAGE.			
GN	RAGE.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;			
OX	NCBI_TaxID=10090;			
RN	[1]			
FP	SEQUENCE FROM N.A.			
PA	Powen L., Mahairas G., Qin S., Ahearn M.E., Dankers C., Lasky S.,			
HA	Loerz C., Schmitt S., Tipton S., Trautwein R., Zaakre R., Hood L.,			
RT	"Sequence of the mouse major histocompatibility locus class III			
RT	region."			
RL	Submitted (01-1997) to the EMBL/GenBank/DBJ databases.			
DP	EMBL: AF010001; ZABR2007 1;			
DR	InterPro: IPR003598; Ig_C2.			
DR	InterPro: IPR003600; Ig_Like.			
DR	InterPro: IPR003606; Ig_MHC.			
DR	SMART: SM00408, IGG2; 1.			
DR	SMART: SM00416, IG_Like; 1.			
DR	PROSITE: PS00346, IG_MHC, UNKNOWN_1.			
KW	Immunoglobulin domain.			
SU	SEQUENCE 402 AA; 4265 MW; D8FDC5UACDC992 CRO64;			

Query Match: 78.2%, Score 1381, DB 11, Length 402.  
Best local similarity: 79.4%, Pred. No. 5, 4e100.  
Matches: 263, Conservation: 19, Mismatches: 45, Indels: 2, Gaps: 2.  
QY 2 QUNIIARIGEPVLAK:KGAPKKKIVKRWKLNTRTEAKKVLSPQGGWDSVARVLPNGS 61  
D6 24 QUNIIARIGEPVLAK:KGAPKKKIVKRWKLNTRTEAKKVLSPQGGWDSVARVLPNGS 82



```

Db 92 LGDCISRDSYEGNITWYRNGKVLQPLEGAVVIFKKQMDPVLTQTYMTSSLEYKTK-A 150
QY 179 DPRFTFSCSPGIPRRKRLRTAPIDRVWE-VPLFEVOLVVEPCGAVAPGCTVTLIC 237
Db 151 DIQIFPFCSTHYGCSQ--KVNHPGAVFDIYPTQVTLQVLPKNAIKEDUNITLKC 208
QY 238 EVPAQSP-QIHWKDKVPLPSPVLIILPEIGPDQGYSC 279
Db 209 LQCNPTTFEFPFVILGCPGICRSSNTYTLTNVRRNATGNKYC 251

RESULT 4
Q95812 PRELIMINARY: PRT: 646 AA.
AC 095812:
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cell surface glycoprotein PHLI2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
FP SEQUENCE FROM N.A.
RA Gui L., Chang L., Browne P.V., Heibel R.P.;
RT "PHI2 from human umbilical vein endothelial cells.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF084868; AAL7794.1; -.
DR InterPro: IPR003594; Ig_c2.
DR InterPro: IPR003600; Ig_c2.
DR InterPro: IPR003606; Ig_MHC.
DR Pfam: PF00047; Ig; 5.
DR SMART: SM00408; IGC2; 2.
DR SMART: SM00410; IG-like; 2.
KW Immunoglobulin domain.
SQ SEQUENCE 646 AA; 71665 MW; 1B5FB8ADD930738E CRC64;

```

```

Query Match 12.48; Score 219.5; DB 4; Length 646;
Best Local Similarity 23.38; Pred. No. 6e-09;
Matches 95; Conservative 56; Mismatches 134; Indels 123; Gaps 19;

QY 4 ITARIGEPVLKCKGAPKPP-QRLW--KLNTGRTEAWKVLSPG---GGPWDSVARVL 57
Db 36 VEVEVGSTALLKCGLSQSGNLSHVDFSVHKKKRTLIFRVHGGGQSGSEGEYEQRLSIQ 95
QY 58 PNG-SILFLPAGVIGDEGIFRCQAMNRNGKETKS-NYRV--RVYQIPGKPEI---V 105
Db 96 DRGATLALTQVTPQDERIFLCQ-----GKPPRSQVEYRIQLPVYKAPPEPNQVNPGLGIPV 150
QY 106 DSASELTAGVNVKVTGCVSGSYDAGTILSWHLDGKPLVPNEKGVSVKQTRRHPTGLPT 165
Db 151 NSKF-----PPEVATCVGNRGYPICQVITWYKNGRPLKEKNRVHI-QSSQIVSESSGLYT 203
QY 166 LQSELMTVPAGSGDPPPTFS--SFSP-ILPHFHALP-----TAPI---QPPVW---EPVPL- 213
Db 204 LQSLIKALQVLE-EDKDAQFYCELNYRLPSGNHMKESREVTVPFYPTKVMLEVEPVGML 262
QY 214 -----EE----- 215
Db 263 KEGDVEIKCLADGNPFPHFISSEUNPSTREAEETTUNGVLVLEPARKHSRYEQG 322
QY 216 -----VOLVVEPEG-----GAVAP-----GGTVTLTCVTPAQSPQIHWKDKGV 254
Db 323 LDLDITMISLSEPPQELLVNVSDVRVSPAAPERQEGSSLLTTCAESSQULLEFQWLREET 382
QY 255 PLPLPSPVLIILPEIGPDQGYSCVAT-HSSHGPOESRAVYSIILEP 301
Db 383 QGVLEKGPVLQHLDLKREAGGYKCVASVPSIPGLNRTQLNVNVAIFGP 430

RESULT 5
Q95812 PRELIMINARY: PRT: 646 AA.
AC 095812:
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Riq-1 protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

ID Q9BRD9 PRELIMINARY: PRT: 646 AA.
AC Q9BRD9:
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similar to melanoma adhesion molecule.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1
FP SEQUENCE FROM N.A.
RA Straussberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC006329; AAH06329.1; -.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR003600; Ig-like.
DR InterPro: IPR003606; Ig_MHC.
DR Pfam: PF00047; Ig; 5.
DR SMART: SM00408; IGC2; 3.
DR SMART: SM00410; IG-like; 1.
KW Immunoglobulin domain.
SQ SEQUENCE 646 AA; 71608 MW; 1F2C696B64B16635 CRC64;

Query Match 12.48; Score 218.5; DB 4; Length 646;
Best Local Similarity 23.38; Pred. No. 7.2e-09;
Matches 95; Conservative 56; Mismatches 134; Indels 123; Gaps 19;

QY 4 ITARIGEPVLKCKGAPKPP-QRLW--KLNTGRTEAWKVLSPG---GGPWDSVARVL 57
Db 36 VEVEVGSTALLKCGLSQSGNLSHVDFSVHKKKRTLIFRVHGGGQSGSEGEYEQRLSIQ 95
QY 58 PNG-SILFLPAGVIGDEGIFRCQAMNRNGKETKS-NYRV--RVYQIPGKPEI---V 105
Db 96 DRGATLALTQVTPQDERIFLCQ-----GKPPRSQVEYRIQLPVYKAPPEPNQVNPGLGIPV 150
QY 106 DSASELTAGVNVKVTGCVSGSYDAGTILSWHLDGKPLVPNEKGVSVKQTRRHPTGLPT 165
Db 151 NSKE-----PPEVATCVGNRGYPICQVITWYKNGRPLKEKNRVHI-QSSQIVSESSGLYT 203
QY 166 LQSELMTVPAGSGDPPPTFS--SFSP-ILPHFHALP-----TAPI---QPPVW---EPVPL- 213
Db 204 LQSLIKALQVLE-EDKDAQFYCELNYRLPSGNHMKESREVTVPFYPTKVMLEVEPVGML 262
QY 214 -----EE----- 215
Db 263 KEGDVEIKCLADGNPFPHFISSEUNPSTREAEETTUNGVLVLEPARKHSRYEQG 322
QY 216 -----VOLVVEPEG-----GAVAP-----GGTVTLTCVTPAQSPQIHWKDKGV 254
Db 323 LDLDITMISLSEPPQELLVNVSDVRVSPAAPERQEGSSLLTTCAESSQULLEFQWLREET 382
QY 255 PLPLPSPVLIILPEIGPDQGYSCVAT-HSSHGPOESRAVYSIILEP 301
Db 383 QGVLEKGPVLQHLDLKREAGGYKCVASVPSIPGLNRTQLNVNVAIFGP 430

RESULT 6
Q92214 PRELIMINARY: PRT: 1344 AA.
AC Q92214:
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Riq-1 protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```



```

PL Brain_Ros_640-296_387(1994)
CC 1- FUNCTION: CELL ADHESION MOLECULE THAT BINDS TO CD6. INVOLVED IN
CC NEURITE EXTENSION BY NEURONS VIA HETEROPHILIC AND HOMOPHILIC
CC INTERACTIONS. MAY PLAY A ROLE IN THE BINDING OF T AND B CELLS TO
CC ACTIVATED LEUKOCYTES. AS WELL AS IN INTERACTIONS BETWEEN CELLS OF
CC THE NERVOUS SYSTEM. MAY BE INVOLVED IN OSTEOGENESIS. MAY PLAY AN
CC IMPORTANT ROLE IN THE DEVELOPMENT OF A VARIETY OF MESENCHYMAL
CC TISSUES (BY SIMILARITY).
CC 1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC 1- TISSUE SPECIFICITY: STRONGEST EXPRESSION IN THE LUNG, THEN BRAIN,
CC LIVER, AND KIDNEY. PRESENT IN THE SOMATOSENSORY SYSTEM, BASAL
CC GANGLIA, CORTEX, OLFACTORY SYSTEM, AND CIRCUMVENTRICULAR ORGANS.
CC 1- PTM: THE N-TERMINAL IS BLOCKED.
CC 1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 4
CC C2-LIKE AND 2 V-LIKE DOMAINS.
CC EXBL: AB064338; BAA23279.1;
CC EXBL: Y13241; CAA73693.1;
CC EXBL: Y13240; CAA73692.1;
CC HSP: Q13740; IKJC.
CC InterPro: IPR003599; Ig_
CC InterPro: IPR003600; Ig_Like
CC InterPro: IPR003606; Ig_MHC.
CC Pfam: PF00047; Ig_5.
CC SMART: SM00409; Ig_3.
CC SMART: SM00410; Ig_Like_2.
KW Cell adhesion, immunoglobulin domain, Glycosylated, Transmembrane
KW Signal: Antigen.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 583 CD166 ANTIGEN.
FT DOMAIN 28 527 EXTRACELLULAR (POTENTIAL).
FT TRANSFM 528 548 POTENTIAL
FT DOMAIN 549 583 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 36 120 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 150 227 IG-LIKE V-TYPE DOMAIN.
FT DOMAIN 263 320 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 347 399 IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN 428 492 IG-LIKE C2-TYPE DOMAIN.
FT DISULFID 43 113 POTENTIAL.
FT DISULFID 157 220 POTENTIAL.
FT DISULFID 270 313 POTENTIAL.
FT DISULFID 354 392 POTENTIAL.
FT CONFLICT 435 485 POTENTIAL.
FT CONFLICT 339 389 S -> G (IN REF. 2).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 167 167 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 361 361 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 480 480 N-LINKED (GLCNAC. ) (POTENTIAL).
FT CARBOHYD 499 499 N-LINKED (GLCNAC. ) (POTENTIAL).
SO SEQUENCE 583 AA; 65021 MW; D36B73B54F5DF61E CRC64.

Query Match 12.1%; Score 213; DB 11; Length 583;
Best Local Similarity 25.3%; Pred. No. 1.7e-08;
Matches 82; Conservative 50; Mismatches 136; Indels 56; Gaps 13;

QY 4 ITARIPEPLVKKAGAPKPPPPLEKLNLTGRTAEKVKLSPPQ438-----PW 50
DB 31 VNSAYGDTIVMPH-----IVPQNIIM-----GK---WKYKPKTSWFIAPFSTKKSQVY 80
QY 51 DVSAR-----VLPNGSIFLPVAGTQGGIFRCQAMRNCKTKSNY-----RVRYQIP 99
DB 81 DQVPEYKDRLSSENVTLSINNNAKIDSEKPEFVCMV-----TEDVFEAPTIVKVKQIF 134
QY 100 GKPEIVDASGLIAGVPNKVIGCVSGYPAGTISWHLCKPLVINKGVSVKQETRRHP 159
DB 135 SKPEIVNRAAFETPGCKIKLGLTISHSYSPHONIHWPNCKVLQVPVDEVSILPKKEIRP 194
QY 160 ETGLFTQSLQSLMTVPARGSDPPPTFS-----FSPGLPRHIALPTAPIQPPWF-PVPLEE 215
DB 195 GTQLIYNTSSLSLVKTKK-SHQLMPTQCSVIYVYGPS-----GUKTIYSPQALEIVYVPEQ 248

Query Match 11.9%; Score 210.5; DB 13; Length 330;
Best Local Similarity 24.7%; Pred. No. 1.3e-08;
Matches 76; Conservative 45; Mismatches 136; Indels 51; Gaps 11;

QY 9 GPEPLVKCGAPKPPPPLEKLNLTGRTAEKVKLSPPQGGWDSVAVPLPNCISLFLPAV- 67
DB 37 GEPALNCK-AEGRPV-TIENYKGESEVETIKQ-DPK-----SHMLLPSSGLEFLRIV 88
QY 68 -----GIGQSGIFPGGLAMNPNKKEIKSNYKRVYQIPGKPEIVLSASELTACVFN-KVGTG 122
DB 89 HGKSRPDEGVYVGVVAVRNLYLGEAVSHNASLEVAII--RDDFRQNPSPDVMVAVGPPAVMEC 146
QY 123 VSENSYPATYLSWHITGTFPLVPNEKSVAKKEVTPKHPETGLTLQSEIMVTTPARGSDPPP 182
DB 147 QPPKRGHPEPTISWKKDSTPLDQKDEKILIR-----GCKLMIYIRKND--- 189
QY 183 FSCSESLPLPHHAI-TAPIQPR---VMEKVPLEEVULVVEPECCAVAPCTVTLCEV 239
DB 190 -----AKYVVVVTNNVGERESEVAELITVEPFSVKRPSN-ATVVDSDSEFKCEA 240
QY 240 PAQSPFQIHWKPKCVPLP-----LPYSVVLILPILPQGGTGYSCYATPSSHQPSRAV 294
DB 241 RQDPVPTVWPKHKGQEDPKARYEIRHUIHLIKPKVMAGIMSGSYTCVAHMMVSGKASATL 300
QY 245 SISLI-EP 301
DB 301 IVQVVSEP 308

RESULT 10
Q92625
ID Q92626 PRELIMINARY; PPT; 149 AA
AC Q92626
DI 01-FEB-1997 (TREMBLER). 02, Created)

QY 216 VOLVVEPECCAVATCCITVLTCEVPAQVSP QIHWKCVPLPLPISVLLIPILPQGGTQ 274
DB 249 VTIQVLPKNAIKESDNITLQCLGNGNPPEPFYLPQAGETSSNTYTLTDOVRNAT 308
QY 275 GTYSVATHSSHSIGPQESPAVSISI 298
DB 309 GDYKC-----SLIDGRNMAASTTI 327

RESULT 9
Q90242
ID Q90242 PRELIMINARY; PPT; 330 AA.
AC Q90242
DI 01-DEC-2001 (TREMBLER). 19, Created)
DI 01-DEC-2001 (TREMBLER). 19, Last sequence update)
DI 01-MAR-2002 (TREMBLER). 20, Last annotation update)
DE Roundabout1 protein (Fragment).
GN ROBO1.
OS Gallus gallus (Chicken).
OC Archaeopteryx, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
OC Archosauria, Aves, Neognathae, Galliformes, Phasianidae, Phasianinae,
OC Gallus.
OC NCBI_TaxID=9031;
RN 111
RP SEQUENCE FROM N.A.
PX MEGALN-21356016; PubMed 11472852;
RA Vargesson N, Juria V, Messina I, Erskine L, Laufer E.;
RT "Expression patterns of Slit and Robo family members during vertebrate
RT limb development.";
PL Mech. Dev. 106:175-190(2001).
DR EMBL: AF364047; AAK94293.1;
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00406; Ig; 1.
FT NON_TER 1 330
FT NON_TER 330 330
FT SEQUENCE 330 AA; 36745 MW; 0613488F78CBE61 CRC64;

Query Match 11.9%; Score 210.5; DB 13; Length 330;
Best Local Similarity 24.7%; Pred. No. 1.3e-08;
Matches 76; Conservative 45; Mismatches 136; Indels 51; Gaps 11;

QY 9 GPEPLVKCGAPKPPPPLEKLNLTGRTAEKVKLSPPQGGWDSVAVPLPNCISLFLPAV- 67
DB 37 GEPALNCK-AEGRPV-TIENYKGESEVETIKQ-DPK-----SHMLLPSSGLEFLRIV 88
QY 68 -----GIGQSGIFPGGLAMNPNKKEIKSNYKRVYQIPGKPEIVLSASELTACVFN-KVGTG 122
DB 89 HGKSRPDEGVYVGVVAVRNLYLGEAVSHNASLEVAII--RDDFRQNPSPDVMVAVGPPAVMEC 146
QY 123 VSENSYPATYLSWHITGTFPLVPNEKSVAKKEVTPKHPETGLTLQSEIMVTTPARGSDPPP 182
DB 147 QPPKRGHPEPTISWKKDSTPLDQKDEKILIR-----GCKLMIYIRKND--- 189
QY 183 FSCSESLPLPHHAI-TAPIQPR---VMEKVPLEEVULVVEPECCAVAPCTVTLCEV 239
DB 190 -----AKYVVVVTNNVGERESEVAELITVEPFSVKRPSN-ATVVDSDSEFKCEA 240
QY 240 PAQSPFQIHWKPKCVPLP-----LPYSVVLILPILPQGGTGYSCYATPSSHQPSRAV 294
DB 241 RQDPVPTVWPKHKGQEDPKARYEIRHUIHLIKPKVMAGIMSGSYTCVAHMMVSGKASATL 300
QY 245 SISLI-EP 301
DB 301 IVQVVSEP 308

RESULT 10
Q92625
ID Q92626 PRELIMINARY; PPT; 149 AA
AC Q92626
DI 01-FEB-1997 (TREMBLER). 02, Created)

```

```

DI 01 FEB 1997 (Tremblrel. 02, last sequence update)
DE 01 JUN 2002 (Tremblrel. 21, last annotation update)
DE MYELOBLAST KIAA0230 (Fragment)
GN KIAA0240.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN 111
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RA MEDLINE=97191544; PubMed=959781;
RA Napsa T., Seki N., Ishikawa K., Ohira M., Kawabayashi Y., Ohara O.,
RA Tanaka A., Kotani H., Miyajima N., Nomura N.;
RA "Prediction of the coding sequences of unidentified human genes. VI.
RI The coding sequences of 40 new genes (K1AA0201-K1AA0280) deduced by
RI analysis of cDNA clones from cell line KG-1 and brain.";
RL DNA Res. 3: 421-429(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=MELANOMA;
RA MEDLINE=95048383; PubMed=7959781;
RA Weiler S.R., Taylor S.M., Evans R.J., Kan Mitchell J., Mitchell M.S.,
RA Trent J.M.;
RA "Assignment of a human melanoma associated gene M650 (h2S448) to
RI chromosome 2p25.3 by fluorescence in situ hybridization.";
RI Genomics 22: 243-244(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=MELANOMA;
RA Mitchell M.S., Kan Mitchell J., Minov R., Edman C., Evans R.J.;
RA "Identification of a novel melanoma gene (M650) - likely the gene for
RI 111 receptor antagonist, which encodes epitopes recognized by human
RI cytolytic T lymphocytes.";
RI Submitted (Oct. 1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: D86983; BAA13219.1;
DR EMBL: AF200348; AAT06354.1;
DR HSSP: P05164; 1CKP.
DR InterPro: IPR002007; Anni peroxidase.
DR InterPro: IPR003598; Iq_C2.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR Clerm.
DR InterPro: IPR000372; LRR Nterm.
DR InterPro: IPR003591; LRR Typ.
DR InterPro: IPR001007; VWF_C.
DR Pfam: PF00047; Iq_4.
DR Pfam: PF00560; LRR_5.
DR Pfam: PF01464; LRRCT_1.
DR Pfam: PF00094; VWC_1.
DR PRINTS: PR00457; ANTEROXIDASE.
DR SMART: SM00408; IGG2_4.
DR SMART: SM00082; IPRCT_1.
DR SMART: SM00013; IPRNT_1.
DR SMART: SM00469; LRR_TYP_4.
DR SMART: SM00214; VWC_1.
DR PROSITE: PS01208; VWF; UNKNWN 1.
KW Immunoglobulin domain.
FT NON TER
SQ SEQUENCE 1496 AA; 167209 MW; 5909A7069B1ABFE CRC64;

Query Match
Best Local Similarity 24.4%; Score 209; DB 11; Length 1496.
Matches 77; Conservative 41; Mismatches 144; Indels 52; Gaps 11.

QY 2 QNTAPGFEFFVVKRFAKFFKQFEWELNIGTELEAKVLSQSGKFWLSNAPVFNPS 41
DB 11 111 111 111 111 111 111 111 111 111 111 111 111 111 111
466 LNFVILWGESVLEER SAUGHPIPRISW-TRGRIHL----- PWRPNVNIIPSG 413
QY 62 LNFVAVLQDDGIFPRGAMNNKPKPKSNVYVGVYVYVYVYVYVYVYVYVYVYV 121
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
414 LYQNVVQSGSGEYA/SATNINDSVIATAF-----ITVALPQFTVTTPQPPV 460

Query Match
Best Local Similarity 25.4%; Score 209; DB 11; Length 1612.
Matches 78; Conservative 44; Mismatches 133; Indels 52; Gaps 12.

QY 9 GELVILKQGAIPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 67
DB 111 111 111 111 111 111 111 111 111 111 111 111 111 111
43 GEPALINCK AGKRPPLPRPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 94
QY 68 --GTGPPGTPPGVAMNPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 122
DB 111 111 111 111 111 111 111 111 111 111 111 111 111
95 HGKSKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 152
QY 123 VSEGSYPATLISWELDKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 182
DB 111 111 111 111 111 111 111 111 111 111 111 111 111
133 QFERCHPEPTISWKKKSGSLGKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 195
QY 183 TFCSEFSGLLPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 239
DB 111 111 111 111 111 111 111 111 111 111 111 111 111
136 -----ACKYVVCNMWTERKSVAEILVIRTSFKPKPKPKPKPKPKPKPK 246
QY 240 EAPSPQIHWKKKSVPLP-----LPSPVLILPEAGPQKQYKSVLHSSHGTSRAV 294

```



Db	247	5-DEPVTVMKUC=SELPKSYEIRFOHILKIRKVTADMGSYTCVA-ENNVGKAEASA-	304
QY	295	SISLIEP 301	
Db	305	TLTVQEP 311	
RESULT 12			
QY6N7			
ID	QY6N7	PRELIMINARY:	PRT: 1651 AA.
AC	QY6N7		
DT	01-NOV-1999	(TREMELrel. 12, Created)	
DT	01-NOV-1999	(TREMELrel. 12, Last sequence update)	
DT	01-DEC-2001	(TREMELrel. 19, Last annotation update)	
DE		Roundabout 1.	
CN	ROBO1		
OS	Homo sapiens (Human).		
OC	Eukaryota, Metazoa, Chordata: Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606,		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=98117249, PubMed=9458045;		
RA	Kidd T., Brose K., Mitchell K.J., Fetter R.D., Tessier-Lavigne M.,		
RA	Goodman C.S., Tear G.;		
RT	"Roundabout controls axon crossing of the CNS midline and defines a		
RT	novel subfamily of evolutionarily conserved guidance receptors.";		
RL	Cell 92:205-215 (1998).		
DR	EMBL, AF040990, AAC39575.1,		
DR	HSSP; P56276; ITLK.		
DR	InterPro: IPR003961; FN-III.		
DR	InterPro: IPR003600; Ig_Like.		
DR	InterPro: IPR003006; Ig_MHC.		
DR	InterPro: IPR003596; Ig_V.		
DR	Pfam: PF00041; f03_3.		
DR	Pfam: PF00047; Ig_5.		
DR	SMART: SM00406; FN3; 3		
DR	SMART: SM00410; Ig_Like; 4.		
DR	SEQUENCE 1651 AA; 180929 MW; 9D98CD7CAB73074D CkC64;		
Query Match			
Best Local Similarity	25.4%;	Pred. NO. 1.3e-07;	
Matches	78; Conservative	44; Mismatches 133; Indels 52; Gaps 12;	
QY	9	CEPLVLKCKAPKKPQRLKNTGRTFAWKVLSPOGGPMDSVARVLPGNSLFLPAV-	67
Db	82	GEATLNCK-AGKRTPILEWKGGKVEIDKD-DPR-----SIRMLLPSSLEFLRIV	133
QY	68	-----GTDEGIFPCOAMNRNGKFTKSNYKRVVYQIPCKPEIVDSASELTAGVFN-KVGTC	122
Db	134	HGKKSRLPDEGVYCVAKNVLGCAVSHNASLEVAILE--RDDFRQNP-SDVMVAVGEFAVMEC	191
QY	123	VSEGSYPAGTFLSHULDSKFLVNEK-VSVKELTKRUPFTLSELTASELMTVTFARGSDPPR	182
Db	192	QPPKCHPEPLISNKKKCSPLLUKDFEPIIR-----GCKIM:TVTRKSD... 234	
QY	183	TPSCSFSPGLPRHRLRTAPIQPR---VMEPVPLEVLVWHPGCAVANGGITVITLCEV	239
Db	235	-----ASKXVWV:INNVGSEPESEVAELTVLEPPEVSKRPSNLAVTVDDSAEPKCEA	285
QY	240	PAQSPQIHWKDCVPLP-----LPIVSPVLILPEIGPQDGTYSCVATHSSHGPPQSRV	294
Db	286	RGDVPTVTRKIDUGELPKSPYRIRDDHTLIKIRKVTAGDMGSYTCVA-ENNVGKAEASA-	343
QY	295	SISLIEP 301	
Db	344	TLTVQEP 350	
RESULT 13			
Q24327			
ID	Q24327	PRELIMINARY:	PRT: 1033 AA.
AC	Q24327		
DT	01-NOV-1996	(TREMELrel. 01, Created)	
DT	01-NOV-1996	(TREMELrel. 01, Last sequence update)	
DT	01-MAR-2002	(TREMELrel. 20, Last annotation update)	
DE		GP160-DTRK precursor.	
CN	CTK op DTRK op CG8067.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OX	Phyloidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RP	SEQUENCE FROM N.A.		
RA	STRAIN=CANTON S;		
RA	MEDLINE=92164624; PubMed=1371458;		
RA	Pulido D., Campuzano S., Koda T., Modotell J., Barbacid M.;		
RA	"Dmrk, a Drosophila gene related to the trk family of neurotrophin		
RA	receptors, encodes a novel class of neural cell adhesion molecule.";		
RL	EMBO J. 11:391-404 (1992).		
DR	EMBL; X63453; CAA45053.1;		
DR	HSSP; P12931; IPWK.		
DR	FlyBase; FBgn0004839; otk.		
DR	InterPro: IPR000719; Euk_pkinase.		
DR	InterPro: IPR003598; Ig_C2.		
DR	InterPro: IPR003600; Ig_Like.		
DR	InterPro: IPR003006; Ig_MHC.		
DR	InterPro: IPR001245; Tyr_pkinase.		
DR	Pfam: PF00047; Ig_5.		
DR	Pfam: PF00069; pk_nase; 1		
DR	ProDom: PD000001; Euk_pkinase; 1.		
DR	SMART: SM00408; IgC2; 3.		
DR	SMART: SM00410; Ig_Like; 2.		
DR	SMART: SM00219; TyKc; 1.		
DR	PPSITE; PS00117; PROTEIN_KINASE_DOM; 1.		
DR	PPSITE; PS00119; PROTEIN_KINASE_TYR; 1		
KW	ATP-binding; Cell adhesion; Immunoglobulin domain; Signal;		
KW	Transferase; Tyrosine-protein kinase.		
FT	SIGNAL 1 23 POTENTIAL		
FT	CHAIN 1 23 Gp160-DTRK		
DR	SEQUENCE 1033 AA; 114367 MW; D282EFCB28ACA8D0 CRC64;		
Query Match			
Best Local Similarity	11.8%;	Score 308; DB 5; Length 1033;	
Matches	88; Conservative	58; Mismatches 139; Indels 108; Gaps 18;	
QY	12	LVLKCC--KGAP-KKPPQRLKNTGRTFAWKVLSPOGGPMDSVARVLPGNSLFLPAVG	68
Db	133	LLLKHVEGASGDLPLEIEWRNSEKLSWKVQ-----LDQHPLIITQPG	179
QY	69	IODEGIFPCOAMNRNGK-ETKSNYKRVVYQ-----IPCKP-----	102
Db	180	SEDDGLYKCTASNAAGRVMSKQGY---VQSSVKCLPR:PPKKNQKMMESWDKQTFICRG	236
QY	103	-----EIVDSASELTAGVFNKVG-TCVSEGSYPAGT-----LSHILD	138
Db	237	KRGCAAGLEALPAAPDELRIVQGP:IGQSIIKEGHTALTCLYELP:FKNCP:QLRWKKD	296
QY	139	CKPLVPNKGVN-----VKEQTR--RHPETGIFTICSHLMVTTPARGGDP	181
Db	297	SKLLPQVLEGGASPPGHSFGSKDALLPEIAPLVHLKQNG--TLSPASIASDAG----	350
QY	182	PTFSCSFSPGLPRHRLRTAPIQPRWEPVLEFVQVVEFGAVAPGGTIVT-ITCEVP	240
Db	351	-OYOCOLO--LEAHAPINSSPGILEV-----IEQLKFPQPTSKNLELDVAVKVCRAQ	402
QY	241	AUFSFQIHWKDCVPLP-----PSPVLILPELSPQDGTYSCVATHSSHGPPQSRV	294
Db	403	GTPTTQVCWVRDSENTLLPDHVEVDANGTLIFRNVNSBHRGNYTCLATNSQGINAIVAI	462
QY	295	SISLIEPCEGPTAGSVGSGGLGTALALGILG	327
Db	463	NV-VVTPKFSVPVPGFIETSEQGTVMVHCQAIG	494



Fri May 30 17:16:40 2003

us-09-872-185b-2.rspt

Page 9

Search completed: May 30, 2003, 15:57:30  
Job time : 79.018 secs

---

